THE CHEMICAL AND ENZYMATIC SYNTHESIS OF POLY(ADP-RIBOSE)

BY

MICHAEL J. LAMBRECHT

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Paul J. Hergenrother, Chair
Professor Scott K. Silverman
Professor Douglas A. Mitchell
Professor David Sarlah

ABSTRACT

Poly(ADP-ribose) synthesis and degradation are important cellular processes associated with the DNA damage response and many other pathways. Traditionally, these processes have been challenging to study due to the heterogeneity associated with the synthesis of poly(ADP-ribose). Discussed here are methods for the synthesis of homogeneous poly(ADP-ribose) as well as the biology that is enabled by these synthetic efforts.

ACKNOWLEDGMENTS

I want to thank my advisor Paul Hergenrother for his critical support and guidance throughout my time here at UIUC. Paul's mentorship has greatly influenced me both scientifically and personally. I would like to thank the members of my committee including Prof. Scott Silverman, Prof. Doug Mitchell, and Prof. David Sarlah. They have truly been a wonderful committee and supported me as they might for their own students. The Chemistry Biology Interface (CBI) Training Plan has been a wonderful part of my graduate experience. I am grateful to many but especially Prof. Wilfred van der Donk in making CBI such an outstanding program. I would like to thank my friend and colleague Dr. Matt Brichacek for his important role in our project as well as my scientific development. I also want to thank my coworkers especially Dr. Rahul Palchuldhuri and Rachel Botham. Finally, I would like to thank my family and loved ones for all their support.

TABLE OF CONTENTS

Chapter 1. Poly(ADP-ribose): Opportunities and Challenges

1.1 Introduction to Poly(ADP-ribose) Synthesis and Degradation

Poly(ADP-ribosyl)ation is a common post-translational modification mediated by the poly(ADP-ribose) polymerase (PARP) protein family and the enzyme poly(ADP-ribose) glycohydrolase (PARG). A family of 17 PARP enzymes exists with 12 of these enzymes possessing mono ADP-ribosyl transferase activity and five possessing the ability to synthesize poly(ADP-ribose) (PAR)[1]. PARP-1 is the most ubiquity expressed of the PARP family and is responsible for the majority of PAR synthesis[2]. When cells undergo DNA damage, PARP-1 associates to the damaged sites and begins the synthesis of PAR. In this process, PARP-1 converts the molecule β-NAD^+ to polymers of PAR, modifying both itself as well as other acceptor proteins (Figure 1.1). The polymers act as a scaffold to recruit a host of non-covalent, PAR-binding, DNA-damage repair proteins such as X-ray cross complementing protein-1 (XRCC1) (Figure 1.1). Ultimately, recruitment of these proteins results in repair of the damaged DNA. Upon successful DNA repair, the glycohydrolase enzyme (PARG) cleaves the polymers to ultimately form monomeric ADP-ribose[3] (Figure 1.1). In addition to PARG, other enzymes are responsible for PAR processing including ARH3, which is the primary enzyme for PAR processing in the mitochondria of the cell[4]. Though much less active toward PAR than PARG, ARH3's role in PAR metabolism is still being elucidated and is perhaps greater than initially thought[5]. Also, PARG and ARH3 are unable to remove the final ADP-ribose unit from acceptor proteins, and several enzymes (MacroD1, MacroD2 and TARG1) exist that possess this capability[6,7].

Enzymatic degradation of PAR chains is an important process, as long chain PAR can trigger cell death through a process known as parthanatos in the event of extensive DNA damage or PARG deficiency[8-11]. While PAR's role in the DNA damage response has been the most well characterized and delineated, it is a post-translational modification with roles in other cellular processes such as replication, transcription, and chromatin regulation[12]. As a result, the list of PAR-binding proteins and protein domains is ever-growing; PAR has been shown or predicted to interact non-covalently with over 500 proteins through at least four distinct protein-binding domains[13-17].

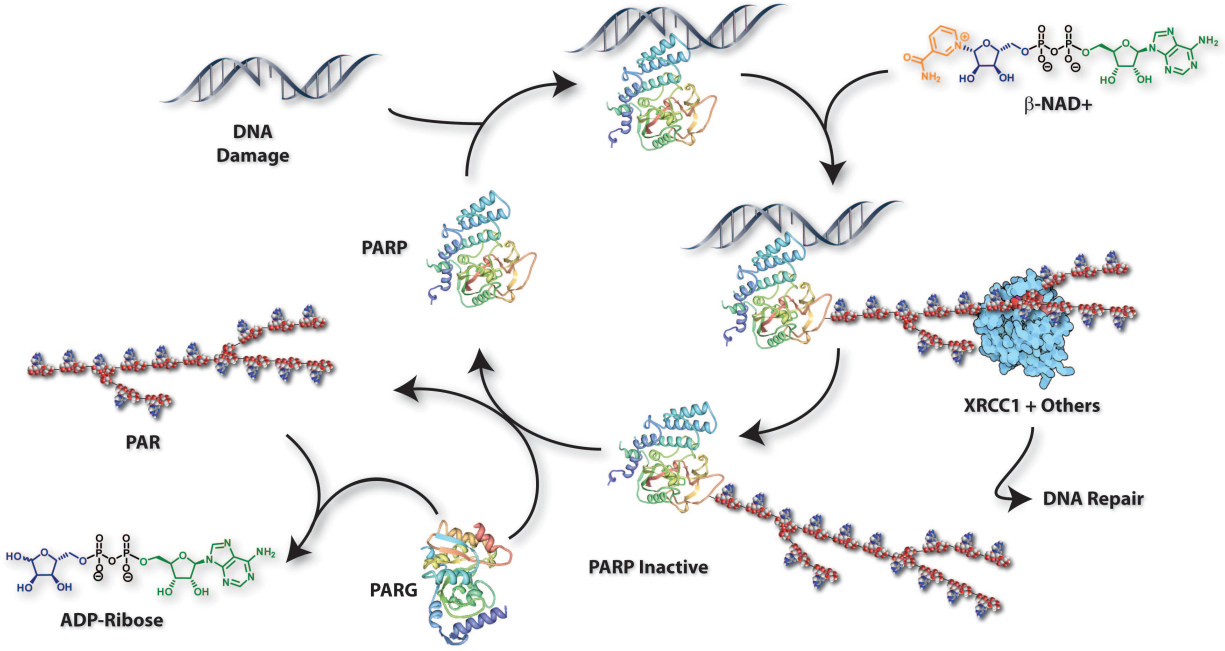


Figure 1.1. Poly(ADP-ribose) synthesis and degrdation in response to DNA damage. PARP: poly(ADP-ribose) polymerase. PARG: poly(ADP-ribose) glycohydrolase. XRCC1: X-ray cross complementing protein 1.

Chemically, PAR is a highly heterogeneous and polydisperse polymer of ADP-ribose with units of ADP-ribose linked via an α-glycosidic linkage between the 2' position of adenosine and the 1'' position of ribose (Figure 1.2). The polymer can be 2 to greater than 200 units in length. The two negative charges per monomeric unit results in polymers with high negative

charge, an aspect believed to be an important factor in the mediation of PAR-protein interactions[18]. A branching process in the polymer occurs with some regularity (about every 20-30 residues), forming an α-glycosidic linkage between the ribose 2'' and 1''' positions[19]. Interestingly, both the polymerization process of branched PAR and the catabolism process to degrade it are mediated by the same enzymes (PARP and PARG) that modulate the synthesis and processing of linear PAR. That is, enzymes that specifically carry out the formation or degradation of branched PAR do not exist.

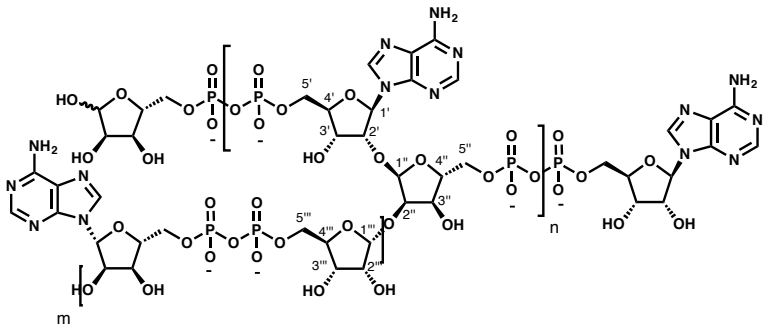


Figure 1.2. Chemical structure of poly(ADP-ribose)

Therapeutic targeting of the PAR cycle has been of interest for several decades[20,21]. Inhibitors of PARP have been pursued extensively as an anti-cancer strategy and have made good progress in the clinic, culminating in the recent approval of Olaparib for advanced ovarian cancers with BRCA mutations[22] as well as castration resistant prostate cancer[23]. A number of clinical trials are currently ongoing pursuing PARP inhibition both as a single agent (primarily in BRCA deficient patients) as well as in combination with DNA damaging agents. Targeting of PARG in the treatment of cancers is another promising approach. However, despite being pursued for some time, PARG inhibition is much less advanced than inhibition of PARP. While good *in vitro* inhibitors of PARG such as ADP-HPD[24-26] and the RBPIs[27] have been discovered, to date no compounds that specifically inhibit PARG activity in whole cells have been described.

1.2 Heterogeneity of PAR as Produced by PARP

The heterogeneity of PAR when produced by PARP has been an impediment to the overall study of PAR metabolism. As previously mentioned, PAR is a heterogeneous and branched polymer produced in lengths ranging from 2 to over 200 monomeric units. PAR is most frequently used as a heterogeneous mixture for *in vitro* studies, and is sometimes fractionated down to narrow distributions by ion exchange chromatography for more exact work. The challenges with accessing homogeneous PAR have been lamented[1] especially in the structural biology literature where co-crystallography of PAR oligomers and polymers with binding proteins is desired[28-30]. Unfortunately, PAR produced *in vitro* is generally long (2-200 units, with the bulk of the distribution being >60), though studies show that the PAR produced in whole cells is often shorter (<20 units)[31]. Therefore, PAR accessed via *in vitro* synthesis may not always be the most relevant for study of biological phenotypes.

1.3 Methods to Address PAR Heterogeneity

1.3.1 Isolation of PAR by Enzymatic Synthesis and Fractionation

PAR oligomers have long been isolated from synthesis by PARP-1 and fractionation by HPLC to provide small amounts of polymers of narrow distribution in length. The use of HPLC for analysis and isolation of PAR synthesized *in vitro* by PARP-1 was first described by Hakam and Kun[32] and further demonstrated by Jacobson[33]. These general procedures have been repeated over the years with very little modification to isolate PAR for use in many studies[9,34-36]. However, the process of PAR production by PARP-1 and fractionation rarely provides polymers of homogeneous length in large quantities for two main reasons. First, expression and

purification of large amounts of PARP-1 is challenging[37]. Second, isolation of polymers of homogenous length by anion exchange chromatography after their synthesis by PARP-1 is a tedious process.

The Mitchison lab attempted to solve the problem of PARP-1 scarcity by using tankyrase, another enzyme with PARP activity, instead of PARP-1 for the synthesis of PAR[37]. This important work showed that tankyrase is easier to express and can be obtained in about 45 fold higher yield than PARP-1. Though tankyrase was shown to be about six fold less active towards PAR synthesis than PARP-1, the higher yields of protein expression more than made up for this difference. Mitchison and coworkers showed that relatively homogeneous PAR oligomers and polymers can be obtained in greater yields using this method than with PARP-1. Others have used this method to obtain the first co-crystal structure of PAR oligomers and a protein[38] as well as to understand the binding profile of XRCC-1 to PAR of different lengths[39]. However, one major drawback of this method is that it does not solve the problem of polymer fractionation following synthesis and this remains the major bottleneck in the process. For instance, while the Mitchison group reported isolation of over 400 mg of unfractionated "bulk" PAR, they were only able to fractionate this amount in 3 mg increments. Therefore, to get meaningful amounts of a polymer of a given length, the purification process needs to be repeated many times. Also, HPLC methods generally do not allow for baseline resolution of polymers of all sizes. Generally, one must design separations of PAR oligomers to get sufficient resolution of a certain distribution of PAR sizes. Resolution of long polymers is rarely achieved.

Control or manipulation of enzymatic synthesis of PAR to selectively provide polymers of defined length has largely been unsuccessful. Varying experimental factors in enzymatic synthesis can have an effect on polymer length, where factors such as substrate concentration,

reaction time, and reaction temperature can all bias formation of shorter or longer polymers. For instance, the Mitchison work demonstrated that by extending reaction times using tankyrase-1, higher amounts of longer polymers could be generated than were present at short times[37]. Still, even at short reaction times, about 50% of the PAR produced is >34 units in length. The addition of an *in trans* acceptor (histones) biased this distribution very little. Also, the Ahel lab has described PARP from the organism *H. aurantiacus*, which produces PAR of shorter lengths (<15 units)[40]. Still, these processes have not been successful in producing homogeneous PAR of defined length and at best provide distributions. To perform a controlled enzymatic synthesis of PAR, one would need to strike a delicate balance between attenuation of PARP activity to prevent runaway polymerization, while maintaining sufficient enzymatic activity so as to provide meaningful amounts of product.

Additionally, one could envision a controlled degradation process where enzymatically synthesized PAR polymers are treated with PARG to afford predominantly short polymers (<10 units). While this approach has not been successfully described in the literature, small amounts of ADP-ribose oligomers (2-4 units in length) have been seen by LCMS after treatment of PAR with human PARG in some studies[38]. Potentially, optimization of this process could enable polymers of a desired length. This approach would face the same hurdles as controlled polymerization (attenuating enzyme activity to yield short oligomers without complete degradation, but retaining enough enzymatic activity such that processing of oligomers takes place).

1.3.2 Chemical Synthesis as a Solution to PAR Heterogeneity

An orthogonal approach to enzymatic synthesis of PAR oligomers is accessing such compounds by total chemical synthesis. While chemical synthesis of PAR could be a solution to the issue of PAR heterogeneity, it is not without its challenges. First, the structure of PAR requires the construction of an α-glycosidic ribofuranose linkage, a major challenge in carbohydrate chemistry[41]. Second, and perhaps more challenging, is the presence of the multiple pyrophosphates in the PAR structure. Construction of multiple pyrophosphates in the same molecule was essentially unknown when we began the project detailed in chapter 2. The chemical synthesis of pyrophosphates is synthetically challenging with most literature conditions involving phosphorimidiazolides or phosphomorphilodates (compounds in the P(V) oxidation state) as activating groups[42] (Figure 1.3a). These methods often involve long reaction times and low yields. Use of the P(III) oxidation state for pyrophosphate formation (Figure 1.3b) was only described in a very few cases[43] when we began the project in chapter 2, but this method has more recently been shown to work with some advantages over the P(V) phosphorimidiazolide and phosphomorphilodate couplings[44,45]. Although a single group has used P(III) pyrophosphites (Figure 1.3c) as a method for pyrophosphate formation[46], we and others[44] have found this method to be irreproducible.

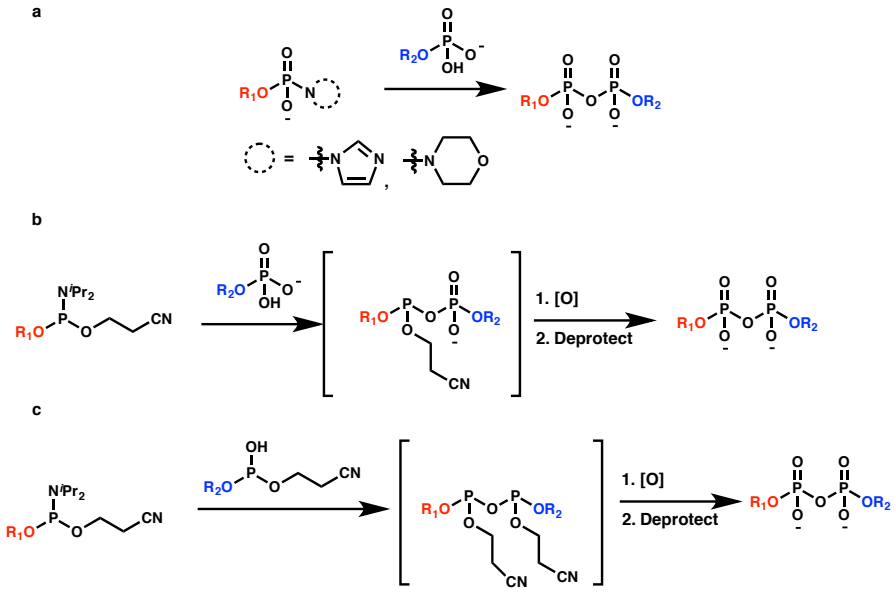


Figure 1.3. Methods for pyrophosphate formation. a. coupling between compounds of the P(V) oxidation state using the common imidazolide or morphilodate methods. b. P(V)-P(III) methods[43] to overcome long reaction times and low yields of a). c. P(III)-P(III) methods described by Parang, *et al*[46]. We and others[44] have found the method in Figure 1.3c not to work as described.

A variety of approaches towards the chemical synthesis of PAR oligomers have been published. In 2008, Mikhailov and coworkers[47] published a synthesis of the PAR disaccharide (Figure 1.4) involving a key glycosylation between an adenosine protected with the Markiewicz reagent[48] (**1-1**) and a protected arabinofuranose (**1-2**). This strategy uses neighboring group participation of the arabinose benzoyl group to efficiently form disaccharide **1-3**. Unfortunately, the use of the arabinose glycosyl donor (**1-2**) to give efficient glycosylation comes with the drawback of the necessity of substantial adjustment after the glycosylation to provide the correct stereochemistry at the c-2'' alcohol. After freeing this position for reaction by debenzoylation and Markiewicz reagent installation, the inversion was performed by a two step oxidation and reduction sequence to give **1-4**. A price was also paid for this procedure, as the use of the Markiewicz reagent to protect both the adenosine 5' and ribose 5'' alcohols means that a strategy for further elaboration of compounds **1-4** or **1-5** to oligomers of PAR in a controlled manner is

needed. Elaboration would require the differentiation of the 5' and 5'' primary alcohols and a strategy for doing this is not obvious.

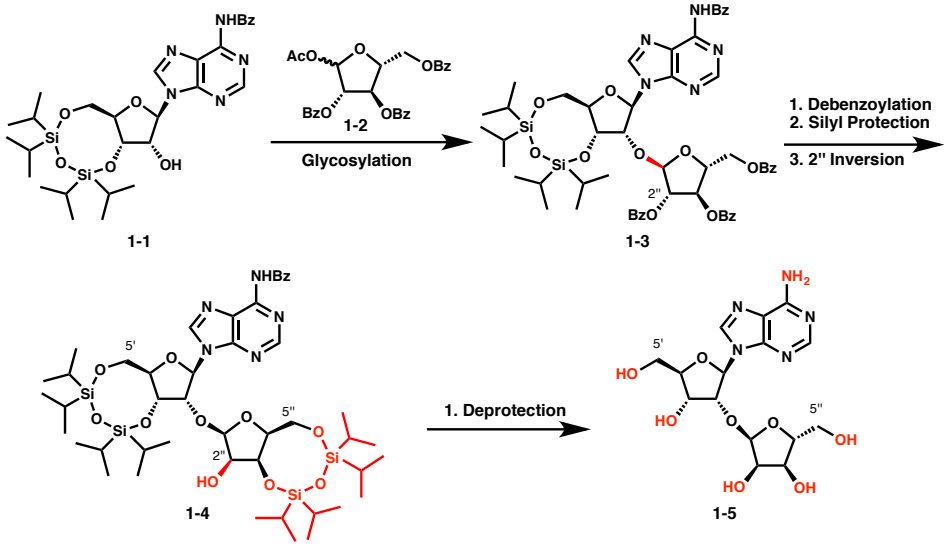


Figure 1.4. Synthesis of the PAR disaccharide (**1-5**) by Mikhailov and coworkers[47].

In an effort to overcome the orthogonality problems of the Mikhailov route, Filippov and coworkers disclosed a route to an orthogonally protected disaccharide to be used as a building block for PAR synthesis[49] (Figure 1.5). Their route relied on the synthesis of an adenosine protected by the Markiewicz reagent as the glycosyl acceptor (**1-6**) and the imidate of tribenzyl ribose as a glycosyl donor (**1-7**). Both the glycosyl donor and glycosyl acceptor can be synthesized in one or two steps from commercial materials. This route provided disaccharide **1-8** in reasonable yields upon glycosylation. Unfortunately, and analogous to the Mikhailov route, the use of the Markiewicz reagent and tribenzyl ribose results in the necessity of a number of post-glycosylation protecting group manipulations. In this case, three steps are required to establish the orthogonality of the 5'' position to give **1-9**, and an additional four are needed to establish full orthogonality in **1-10**. Installation of the requisite phosphorous functionality for fragment coupling reactions from **1-10** would likely require four additional steps. This route does

demonstrate the feasibility of and provide precedent for an α-selective glycosylation to form the desired disaccharide, though it also details the challenges of doing so in light of substantial N-glycosylation and poor reactivity depending on the identity of the glycosyl donor as reported by the authors.

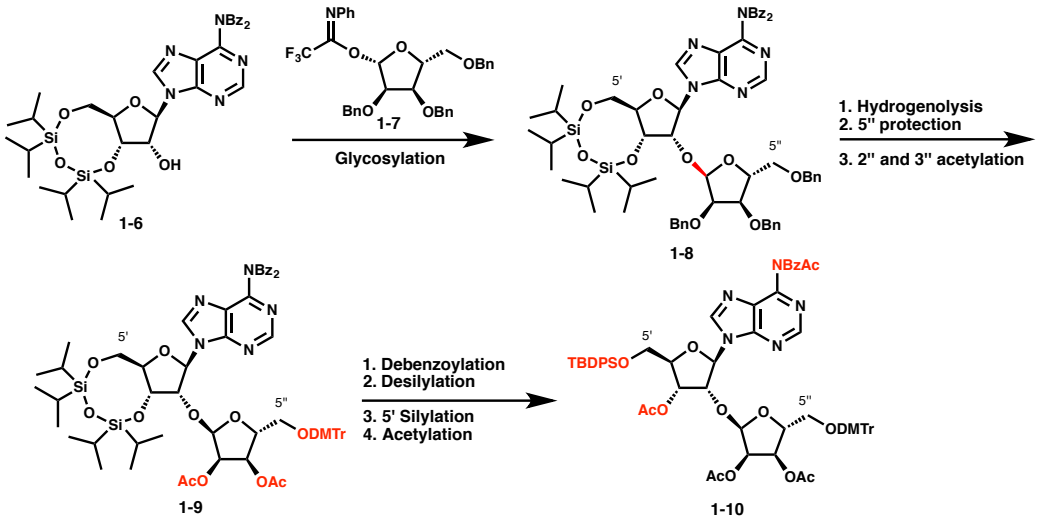


Figure 1.5. Filippov's initial approach to an orthogonally protected disaccharide (**1-10**) for PAR synthesis[49].

Simultaneous to our publication of the material detailed in chapters 2 and 3, Filippov and coworkers disclosed the synthesis of unnatural ADP-ribose, as well as unnatural ADP-ribose dimers and trimers containing a methoxy acetal group on the reducing sugar[45]. This route was enabled by a new synthetic approach to the iterative building block involving a clever Vorbrüggen glycosylation[50] for installation of the nucleobase (Figure 1.6). Glycosyl donor **1-11** was synthesized using a known procedure[51] and glycosyl acceptor **1-12** is commercially available. Formation of the ribose-ribose glycosidic linkage proceeded smoothly to provide disaccharide **1-13**. Vorbrüggen glycosylation with N^6-benzoyladenine (**1-14**) installed the nucleobase of disaccharide **1-15**. A series of five steps provided disaccharide **1-16** with the requisite phosphorous functionality for fragment couplings.

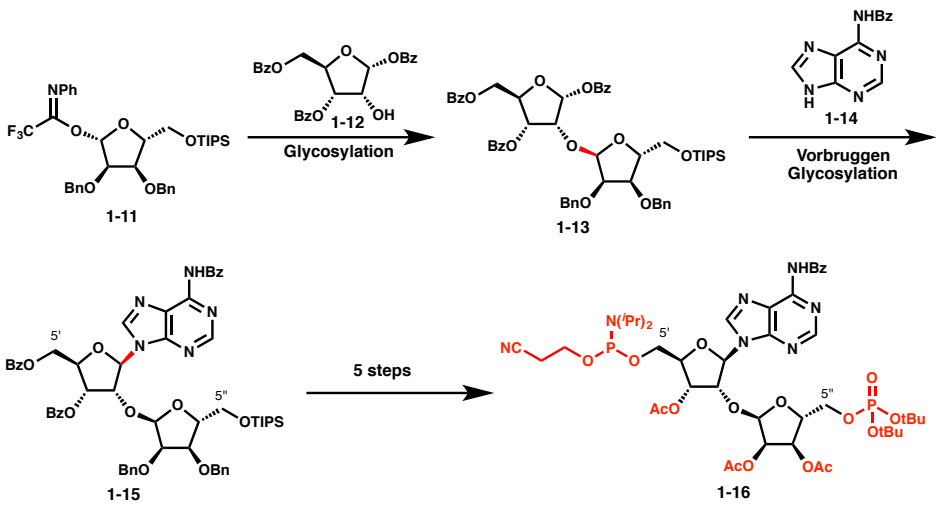


Figure 1.6. Second-generation synthesis of a disaccharide (**1-16**) for fragment couplings by the Filippov group[45].

Formation of the unnatural ADP-ribose dimer and trimer by Filippov and coworkers was performed using a solid phase approach (Figure 1.7). Coupling between ribose phosphate **1-17** and phosphoramidite **1-18** provided O-methoxy ADP-ribose (**1-19**) after deprotection from the resin. A first coupling of disaccharide building block **1-16** in place of adenosine phosphoramidite **1-18** followed by phosphate deprotection and a second coupling with adenosine **1-18** provided O-methoxy ADP-ribose dimer (**1-20**). The analogous trimer (**1-21**) could be synthesized in a similar manner by two couplings with **1-16** and a single one with **1-18**. This approach has the key advantage that the couplings are performed on the solid phase and therefore the purification of intermediates is relatively straightforward. The drawbacks are that the deprotection of the methoxy acetal of compounds **1-19**, **1-20**, and **1-21** to give natural oligomers was not demonstrated in the work, likely because of the challenge of such a reaction. Acetal containing oligomers, such as **1-19**, **1-20**, and **1-21**, are not found in any biological source. The choice of this group was likely due to synthetic ease, and protection of this position with the proper group that would allow for subsequent removal was a major challenge in our work as discussed in

chapter 2. The strongly acidic conditions to affect phosphate deprotection in this work coupled with the acid-lability of anomeric substituents likely precludes the use of many potential protecting group for this position and may explain the choice of the acetal. Additionally, if unnatural oligomers are to be synthesized, the most logical and enabling choice for the unnatural position might be a functionality that allows for probe synthesis, such as the alkyne we synthesized in chapter 2. The methoxy group, however, provides no such advantage. Finally, the stoichiometry of the coupling reactions employed by Filippov require three equivalents of the synthetically most challenging substrate (**1-16**) for each coupling.

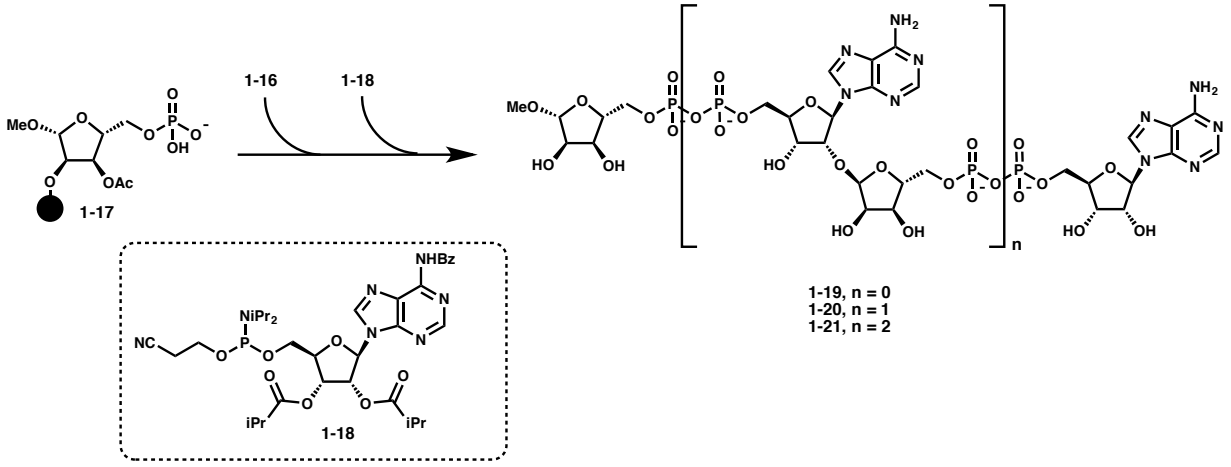


Figure 1.7. Solid phase fragment couplings as described by Filippov and coworkers[45].

The route described by Filippov and coworkers[45] does lend itself well to the synthesis of the PAR trisaccharide of the branched polymer and they have disclosed the synthesis of this molecule[52] (Figure 1.8). This route relies on a similar ribose-ribose glycosylation between glycosyl donor **1-7** and glycosyl acceptor **1-12** to form disaccharide **1-22**. A low-yielding two-step protecting group shuffle was required to free the 2'' alcohol for the second glycosylation to form disaccharide **1-23**. A second glycosylation provided the trisaccharide **1-24**, which was converted to compound **1-25** after a four step sequence to juggle protecting groups. Vorbrüggen glycosylation provided **1-26**. In principle, **1-26** could be elaborated to a substrate that could be

used to form branched oligomers containing pyrophosphates, as the work did show the ability to deprotect the TBDPS group in the presence of the TIPS groups. However, selective deprotection of the 5' benzoyl group was not shown and would likely require three additional steps. Selective formation of PAR oligomers containing the branched trisaccharide would also require orthogonal protection of the phosphate groups if oligomers containing different substitution on the 5'' and 5''' groups were desired, though the setup employed could be easily extended to 5'' and 5''' groups containing the same substitution.

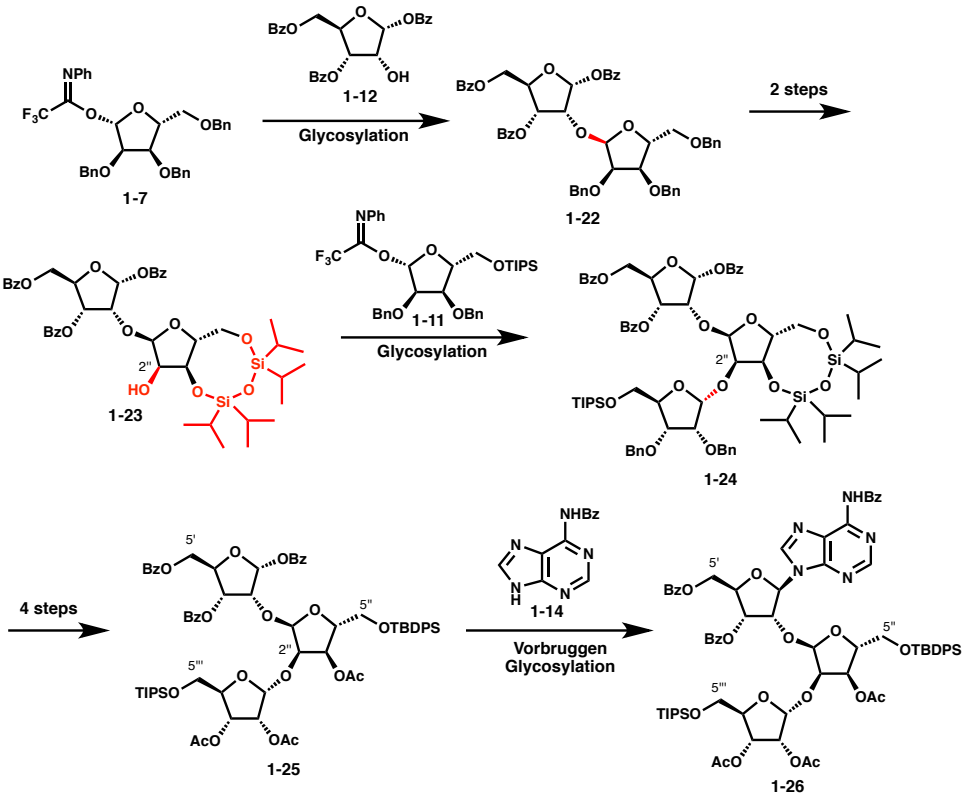


Figure 1.8. The route by Filippov *et al.* to a PAR trisaccharide (**1-26**)[52]

1.3.3 Chemoenzymatic Approaches to PAR Synthesis

In principle, one could imagine a variety of methods through which polymers could be synthesized chemoenzymatically. For instance, the ability of PARP to accept unnatural β-NAD+

substrates has been used for applications mostly involving protein labeling[53,54]. The collection of non-natural β-NAD^+ analogs demonstrated to be PARP-1 substrates are depicted in Figure 1.9. Importantly, one of these approaches describes the ability of PARP-1 to terminate PAR chains by chemical modification of the 2' position of β-NAD^+[54]. Could this termination be performed and subsequently reversed, in principle an approach to homogeneous PAR production and homogeneous PAR polymers could be developed. Efforts towards such an approach are described in chapter 4.

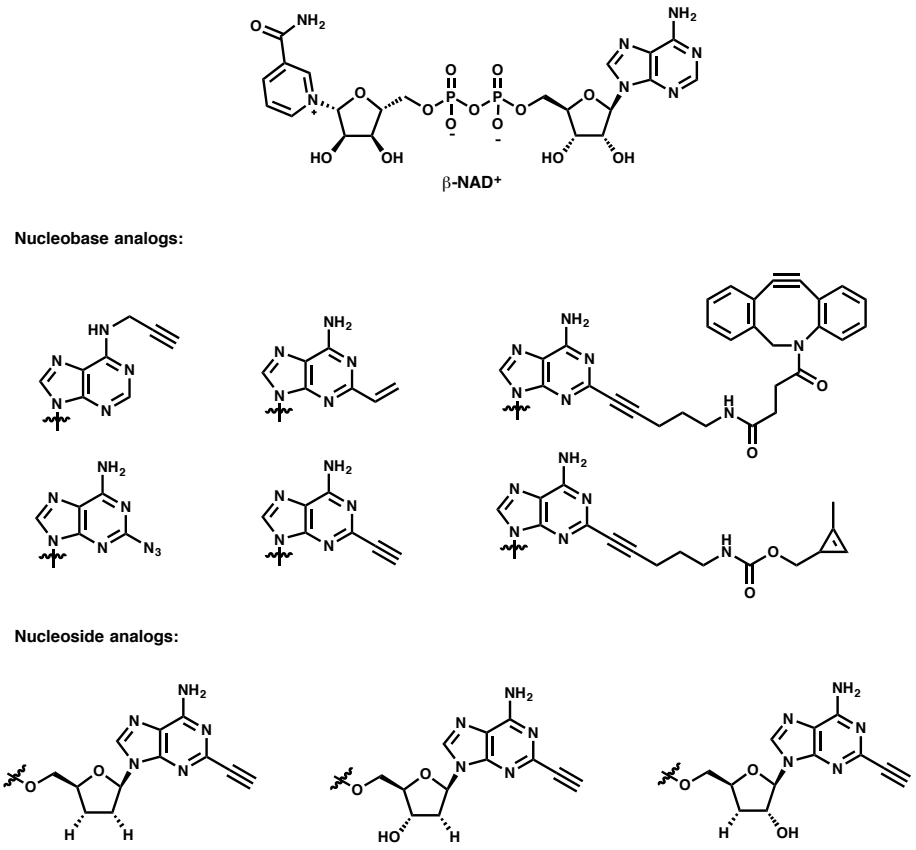


Figure 1.9. β-NAD^+ and analogs that have been shown to be PARP-1 substrates in the literature.

1.4 Conclusions and Outlook

PAR biology has been the subject of intense focus in the past decade, in large part due to the progression of PARP inhibitors, such as Olaparib, through the clinic. While most researchers began study of PARylation due to the therapeutic potential of the process with respect to DNA damage repair, the increased attention to PARylation has uncovered the role of this post-translational modification in a wide range of cellular contexts and processes. Therefore, as knowledge of the role of PAR in the cell continues to grow, the need and desire for tools to understand these processes will as well. The desire for one such tool, homogeneous PAR polymers and oligomers, has been clearly articulated in the literature[1,28-30,55].

To access homogeneous ADP-ribose oligomers, both chemical and enzymatic synthesis approaches have been described. New enzymatic approaches that utilize tankyrase instead of PARP-1 show promise in generation of larger amounts of PAR than was obtainable previously[37]. Unfortunately, these approaches still require tedious fractionation of oligomers and polymers, and therefore do not overcome the major bottleneck in the processes of accessing homogeneous PAR. On the other hand, chemical approaches to PAR synthesis have shown the ability to synthesize short oligomers in greater yield and purity than what is obtainable by enzymatic synthesis. The drawback to such approaches, however, is the number of steps required to access the target compounds. The large number of steps and inefficient fragment coupling reactions associated with them likely precludes the synthesis of long oligomers.

1.5 References

1. Leung, A. K. L. Poly(ADP-ribose): An organizer of cellular architecture. *J. Cell Biol.* **205,** 613–619 (2014).
2. Kraus, W. L. & Hottiger, M. O. PARP-1 and gene regulation: Progress and puzzles. *Mol. Aspects Med.* **34,** 1109–1123 (2013).

3. Hatakeyama, K., Nemoto, Y., Ueda, K. & Hayaishi, O. Purification and characterization of poly(ADP-ribose) glycohydrolase. Different modes of action on large and small poly(ADP-ribose). *J. Biol. Chem.* **261,** 14902–14911 (1986).
4. Niere, M. *et al.* ADP-ribosylhydrolase 3 (ARH3), Not Poly(ADP-ribose) Glycohydrolase (PARG) Isoforms, Is Responsible for Degradation of Mitochondrial Matrix-associated Poly(ADP-ribose). *J. Biol. Chem.* **287,** 16088–16102 (2012).
5. Mashimo, M., Kato, J. & Moss, J. ADP-ribosyl-acceptor hydrolase 3 regulates poly(ADP-ribose) degradation and cell death duringoxidative stress. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 18964–18969 (2013).
6. Žaja, R., Mikoč, A., Barkauskaite, E. & Ahel, I. Molecular Insights into Poly(ADP-ribose) Recognition and Processing. *Biomolecules* **3,** 1–17 (2013).
7. Barkauskaite, E., Jankevicius, G. & Ahel, I. Structures and Mechanisms of Enzymes Employed in the Synthesis and Degradation of PARP-Dependent Protein ADP-Ribosylation. *Mol. Cell* **58,** 935–946 (2015).
8. Koh, D. W. Failure to degrade poly(ADP-ribose) causes increased sensitivity to cytotoxicity and early embryonic lethality. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 17699–17704 (2004).
9. Yu, S.-W. *et al.* Apoptosis-inducing factor mediates poly(ADP-ribose) (PAR) polymer-induced cell death. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 18314–18319 (2006).
10. Andrabi, S. A. S. *et al.* Poly(ADP-ribose) (PAR) polymer is a death signal. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 18308–18313 (2006).
11. Wang, Y. *et al.* Poly(ADP-Ribose) (PAR) Binding to Apoptosis-Inducing Factor Is Critical for PAR Polymerase-1-Dependent Cell Death (Parthanatos). *Sci. Signal.* **4,** ra20–ra20 (2011).
12. Burkle, A. & Virág, L. Poly(ADP-ribose): PARadigms and PARadoxes. *Mol Aspects Med* **34,** 1046–1065 (2013).
13. Krietsch, J. *et al.* Reprogramming cellular events by poly(ADP-ribose)-binding proteins. *Mol Aspects Med* **34,** 1066–1087 (2013).
14. Gagne, J. P. *et al.* Proteome-wide identification of poly(ADP-ribose) binding proteins and poly(ADP-ribose)-associated protein complexes. *Nucleic Acids Res.* **36,** 6959–6976 (2008).
15. Gagne, J. P. *et al.* Quantitative proteomics profiling of the poly(ADP-ribose)-related response to genotoxic stress. *Nucleic Acids Res.* **40,** 7788–7805 (2012).
16. Ahel, I. *et al.* Poly(ADP-ribose)-binding zinc finger motifs in DNA repair/checkpoint proteins. *Nature* **451,** 81–85 (2008).
17. Pleschke, J. M., Kleczkowska, H. E., Strohm, M. & Althaus, F. R. Poly(ADP-ribose) binds to specific domains in DNA damage checkpoint proteins. *J. Biol. Chem.* **275,** 40974–40980 (2000).
18. Ferro, A. M. & Olivera, B. M. Poly(ADP-ribosylation) in vitro. Reaction parameters and enzyme mechanism. *J. Biol. Chem.* **257,** 7808–7813 (1982).
19. Alvarez-Gonzalez, R. & Jacobson, M. K. Characterization of polymers of adenosine diphosphate ribose generated in vitro and in vivo. *Biochemistry* **26,** 3218–3224 (1987).
20. Rouleau, M., Patel, A., Hendzel, M. J., Kaufmann, S. H. & Poirier, G. G. PARP inhibition: PARP1 and beyond. *Nat. Rev. Cancer* **10,** 293–301 (2010).
21. Curtin, N. J. Poly(ADP-ribose) polymerase (PARP) and PARP inhibitors. *Drug Discov. Today Dis. Models* **9,** e51–e58 (2012).

22. Release, F. N. FDA approves Lynparza to treat advanced ovarian cancer. (2014).
23. Release, A. P. lynparza™ (olaparib) granted breakthrough therapy designation by us fda for treatment of brca1/2 or atm gene mutated metastatic castration resistant prostate cancer. (2016).
24. Slama, J. T., Aboul-Ela, N. & Jacobson, M. K. Mechanism of inhibition of poly(ADP-ribose) glycohydrolase by adenosine diphosphate (hydroxymethyl)pyrrolidinediol. *J. Med. Chem.* **38,** 4332–4336 (1995).
25. Slama, J. T. *et al.* Specific Inhibition of Poly(ADP-ribose) Glycohydrolase by Adenosine Diphosphate (Hydroxymethyl)pyrrolidinediol. *J. Med. Chem.* **38,** 389–393 (1995).
26. Koh, D. W. *et al.* SAR analysis of adenosine diphosphate (hydroxymethyl)pyrrolidinediol inhibition of poly(ADP-ribose) glycohydrolase. *J. Med. Chem.* **46,** 4322–4332 (2003).
27. Finch, K. E., Knezevic, C. E., Nottbohm, A. C., Partlow, K. C. & Hergenrother, P. J. Selective Small Molecule Inhibition of Poly(ADP-Ribose) Glycohydrolase (PARG). *ACS Chem. Biol.* **7,** 563–570 (2012).
28. Oberoi, J. *et al.* Structural Basis of Poly(ADP-ribose) Recognition by the Multizinc Binding Domain of Checkpoint with Forkhead-associated and RING Domains (CHFR). *J. Biol. Chem.* **285,** 39348–39358 (2010).
29. Eustermann, S. *et al.* Solution structures of the two PBZ domains from human APLF and their interaction with poly(ADP-ribose). *Nat. Struct. Mol. Biol.* **17,** 241–243 (2010).
30. Wang, Z. *et al.* Recognition of the iso-ADP-ribose moiety in poly(ADP-ribose) by WWE domains suggests a general mechanism for poly(ADP-ribosyl)ation-dependent ubiquitination. *Genes Dev.* **26,** 235–240 (2012).
31. Malanga, M. & Althaus, F. R. Poly(ADP-ribose) Molecules Formed during DNA Repair in Vivo. *J. Biol. Chem.* **269,** 17691–17696 (1994).
32. Hakam, A. & Kun, E. High-performance liquid chromatography of in vitro synthesized poly(adp-ribose) on ion-exchange columns, separation of oligomers of varying chain length and estimation of apparent branching. *Journal of Chromatography A* **330,** 287–298 (1985).
33. Kiehlbauch, C. C. C., Aboul-Ela, N., Jacobson, E. L. E., Ringer, D. P. D. & Jacobson, M. K. M. High resolution fractionation and characterization of ADP-ribose polymers. *Anal. Biochem.* **208,** 26–34 (1993).
34. Fahrer, J., Kranaster, R., Altmeyer, M., Marx, A. & Burkle, A. Quantitative analysis of the binding affinity of poly(ADP-ribose) to specific binding proteins as a function of chain length. *Nucleic Acids Res.* **35,** e143–e143 (2007).
35. Fahrer, J. *et al.* High-Affinity Interaction of Poly(ADP-ribose) and the Human DEK Oncoprotein Depends upon Chain Length. *Biochemistry* **49,** 7119–7130 (2010).
36. Popp, O. *et al.* Site-Specific Noncovalent Interaction of the Biopolymer Poly(ADP-ribose) with the Werner Syndrome Protein Regulates Protein Functions. *ACS Chem. Biol.* **8,** 179–188 (2013).
37. Tan, E. S., Krukenberg, K. A. & Mitchison, T. J. Large-scale preparation and characterization of poly(ADP-ribose) and defined length polymers. *Anal. Biochem.* **428,** 126–136 (2012).
38. Barkauskaite, E. *et al.* Visualization of poly(ADP-ribose) bound to PARG reveals inherent balance between exo- and endo-glycohydrolase activities. *Nat. Commun.* **4,** 1–8 (2013).
39. Kim, I.-K., Stegeman, R. A., Brosey, C. A. & Ellenberger, T. A Quantitative Assay Reveals Ligand Specificity of the DNA Scaffold Repair Protein XRCC1 and Efficient

Disassembly of Complexes of XRCC1 and the Poly(ADP-ribose) Polymerase 1by Poly(ADP-ribose) Glycohydrolase. *J. Biol. Chem.* **290,** 3775–3783 (2015).
40. Slade, D. *et al.* The structure and catalytic mechanism of a poly(ADP-ribose) glycohydrolase. *Nature* **477,** 616–620 (2011).
41. Demchenko, A. V. *Handbook of Chemical Glycosylation*. (John Wiley & Sons, 2008).
42. Wagner, G. K., Pesnot, T. & Field, R. A. A survey of chemical methods for sugar-nucleotide synthesis. *Nat Prod Rep* **26,** 1172–1194 (2009).
43. Gold, H. *et al.* Synthesis of Sugar Nucleotides by Application of Phosphoramidites. *J. Org. Chem.* **73,** 9458–9460 (2008).
44. Kistemaker, H. A. V., Meeuwenoord, N. J., Overkleeft, H. S., van der Marel, G. A. & Filippov, D. V. On the Synthesis of Oligonucleotides Interconnected through Pyrophosphate Linkages. *Eur. J. Org. Chem* **2015,** 6084–6091 (2015).
45. Kistemaker, H. A. V. *et al.* Synthesis of well-defined adenosine diphosphate ribose oligomers. *Angew. Chem. Int. Ed. Engl.* **54,** 4915–4918 (2015).
46. and, Y. A. & Parang, K. Solid-Phase Synthesis of Symmetrical 5‘,5‘-Dinucleoside Mono-, Di-, Tri-, and Tetraphosphodiesters. *Org. Lett.* **9,** 4483–4486 (2007).
47. Mikhailov, S. N., Kulikova, I. V., Nauwelaerts, K. & Herdewijn, P. Synthesis of 2′-O-α-d-ribofuranosyladenosine, monomeric unit of poly(ADP–ribose). *Tetrahedron* **64,** 2871–2876 (2008).
48. Markiewicz, W. T. & Wiewiorowski, M. A new type of silyl protecting groups in nucleoside chemistry. *Nucleic Acids Res.* **1,** s185–s190 (1978).
49. van der Heden van Noort, G. J., Overkleeft, H. S., van der Marel, G. A. & Filippov, D. V. Ribosylation of Adenosine: An Orthogonally Protected Building Block for the Synthesis of ADP-Ribosyl Oligomers. *Org. Lett.* **13,** 2920–2923 (2011).
50. Niedballa, U. & Vorbrüggen, H. A general synthesis of pyrimidine nucleosides. *Angew. Chem. Int. Ed. Engl.* **9,** 461–462 (1970).
51. Kistemaker, H. A. V., van der Heden van Noort, G. J., Overkleeft, H. S., van der Marel, G. A. & Filippov, D. V. Stereoselective Ribosylation of Amino Acids. *Org. Lett.* **15,** 2306–2309 (2013).
52. Kistemaker, H. A. V., Overkleeft, H. S., van der Marel, G. A. & Filippov, D. V. Branching of poly(ADP-ribose): Synthesis of the Core Motif. *Org. Lett.* **17,** 4328–4331 (2015).
53. Jiang, H., Kim, J. H., Frizzell, K. M., Kraus, W. L. & Lin, H. Clickable NAD Analogues for Labeling Substrate Proteins of Poly(ADP-ribose) Polymerases. *J. Am. Chem. Soc.* **132,** 9363–9372 (2010).
54. Wang, Y., Rösner, D., Grzywa, M. & Marx, A. Chain-Terminating and Clickable NAD^+ Analogues for Labeling the Target Proteins of ADP-Ribosyltransferases. *Angew. Chem. Int. Ed.* **53,** 8159–8162 (2014).
55. Daniels, C. M., Ong, S.-E. & Leung, A. K. L. The Promise of Proteomics for the Study of ADP-Ribosylation. *Mol. Cell* **58,** 911–924 (2015).

Chapter 2. Chemical Synthesis of the Natural and Propargyl ADP-Ribose Dimers

Portions of this chapter are reprinted with permission from: Lambrecht, M. J.; Brichacek, M.; Barkauskaite, E.; Ariza, A.; Ahel, I.; and Hergenrother, P. J. *J. Am. Chem. Soc.* **2015**, *137*, 3558-3564. This work was performed in close collaboration with Dr. Matthew Brichacek.

2.1 Synthesis of the Natural ADP-ribose Dimer

2.1.1 The ADP-ribose Dimer as a Model ADP-ribose Oligomer

ADP-ribose oligomers and polymers are produced upon DNA damage and can have many binding partners in cells (>500 proteins with at least four different PAR binding motifs have been predicted or shown to bind PAR)[1-4]. For some PAR binding proteins, the length of the PAR oligomer is an important factor for binding affinity[5]. However, for other classes of proteins the polymer-binding motif is hypothesized to be a much reduced portion of the polymer[1,6]. Determination of the binding unit of the polymer to the protein in these cases is done with simple PAR fragments (ADP-ribose and iso-ADP-ribose). Generally, if a protein binds to one of these fragments, it is concluded that the portion of the PAR polymer represented by the fragment composes the binding motif of the polymer to the protein. However, such molecules do not retain certain key components of the PAR polymer as a whole (multiple pyrophosphates and the α-glycosidic linkage). Therefore, while useful in many studies, model compounds such as ADP-ribose or iso-ADP-ribose lack the structural complexity and important aspects of the PAR polymer and therefore may not represent PAR with sufficient rigor. Homogeneous PAR

oligomers themselves are of course challenging to access by enzymatic synthesis as detailed in chapter 1.

It was therefore envisioned that the simplest ADP-ribose oligomer, the ADP-ribose dimer, would be of value in biochemical studies where ADP-ribose and iso-ADP-ribose fall short. The ADP-ribose dimer contains all the key structural features of the full polymer (the glycosidic linkage and multiple pyrophosphates), and the molecule has been described in the literature as an "ideal ligand for x-ray co-crystal structure determination."[7] While the ADP-ribose dimer has been detected by a variety of methods including MS analysis[8], this compound has not been isolated and handled as a pure substance due to its low levels of production by enzymatic synthesis[9] and challenging preparative separation and isolation. The ADP-ribose dimer was therefore an appropriate and attractive target for total chemical synthesis. If the ADP-ribose dimer could be accessed by a modular chemical route, the compound could be obtained in quantities and purities that are not currently available by enzymatic synthesis and fractionation. Additionally, its synthesis could serve as a template for the synthesis of longer or unnatural oligomers.

2.1.2 Retrosynthetic Analysis and Strategy

Despite the synthetic challenges associated with the α-selective ribose glycosylation and pyrophosphate formation (as detailed in chapter 1), retrosynthetically these disconnections remained the most logical when considering routes to the ADP-ribose dimer (Figure 2.1). The ADP-ribose dimer (**2-1**) was therefore envisioned to originate from pyrophosphate couplings between protected ribose phosphate **2-2**, orthogonally protected disaccharide **2-3** containing the synthetically challenging α-ribose glycosidic linkage, and protected adenosine monophosphate **2-4**. All these precursors were unknown when this project began, with disaccharide **2-3**

representing the greatest synthetic challenge. While protected adenosine monophosphate **2-4** had not been previously prepared, its synthesis was conceptually similar to that of a known compound[10]. Ribose phosphate **2-2**, while relatively simple, provided a challenge with respect to the selection of the proper protecting group for the anomeric position. It was envisioned that setting the α-glycosidic linkage would be the key step for disaccharide **2-3** and that this linkage could be installed via a 1,2-*cis*-selective glycosylation of an orthogonally protected adenosine (**2-5**) by an appropriately protected ribose (**2-6**). Synthesis of the glycosyl donor (**2-6**) and glycosyl acceptor (**2-5**) would commence from ribose and adenosine, respectively. Strategically, for the synthesis as a whole, it was envisioned that the secondary alcohols would ideally be protected as silyl ethers and the nitrogenous bases would be protected with the benzoyl group. These groups were chosen for their stability as well as their ability to be removed quickly, in good yield, and under mild conditions at the end of the synthesis.

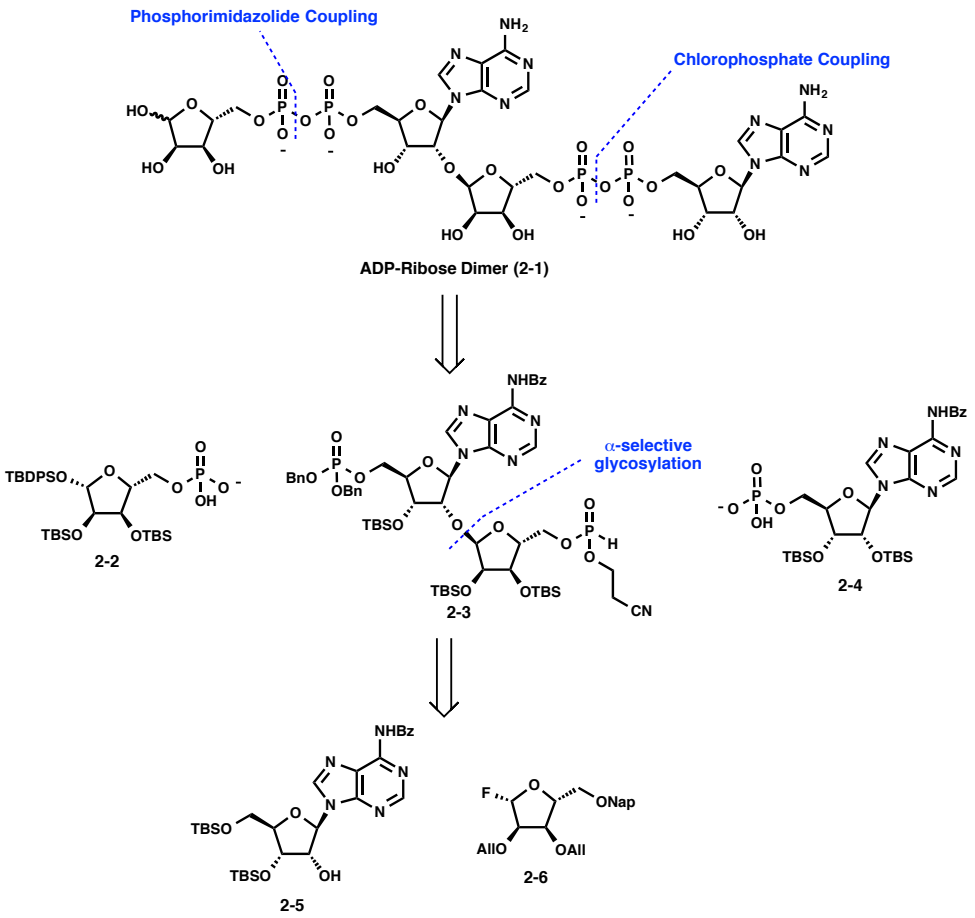


Figure 2.1. Retrosynthetic analysis of natural ADP-ribose dimer (**2-1**).

2.1.3 Selective Synthesis of the Glycosyl Acceptor **2-5**

In the forward sense, the first challenge encountering the synthesis of disaccharide **2-3** was the creation of an appropriately protected adenosine with the 2' alcohol free for chemical glycosylation (such as alcohol **2-5**). While protection of the 3' and 5' alcohols in nucleic acids is commonly carried out using the Markiewicz Reagent[11] to form a cyclic silyl ether, this approach was ruled out due to problems in selectively deprotecting the 5' alcohol in the presence of the 3' alcohol. Therefore, a challenge existed in how to selectively silylate the 5' and 3' alcohols without silylation of the 2' alcohol. Traditionally, conditions have existed for selective 2' silylation with a variety of groups though selective 3' protection is less established. Additionally,

we found the yields and selectivities reported for silver promoted 3' silylation[12] to not be robustly reproducible in our hands.

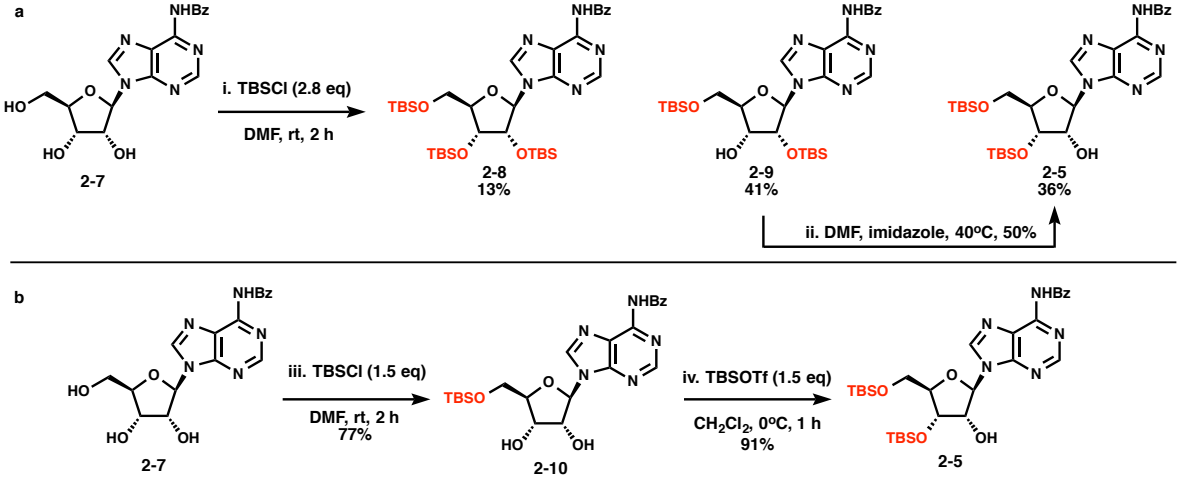


Figure 2.2. a) Unselective synthesis and isomerization of glycosyl acceptor **2-5** b) selective two step synthesis of glycosyl acceptor **2-5** i. TBSCl (2.8 eq), imidazole (6 eq), DMF, 0°C to rt, 1 h, 13% **2-8**, 41% **2-9**, 36% **2-5**. ii. imidazole (2 eq), DMF, 30°C, 50% iii. TBSCl (1.5 eq), imidazole (3 eq), DMF, 0°C, 1 h, 77%. iv. TBSOTf (1.7 eq), THF, -78°C, 1 h, 91% **2-5**, 6% **2-9**.

To preliminarily access such a compound, a known procedure was followed and a non-selective mixture of compounds **2-8**, **2-9**, and the desired **2-5** was obtained (Figure 2.2a)[13]. It had been noted in the literature that isomerization of constitutional isomers **2-9** and **2-5** was possible[14] and indeed this was found to be a reasonable approach to procure more of the desired **2-5**. While this unselective silylation and isomerization approach was useful to provide the alcohol **2-5** in reasonable quantities, scaling the procedure up was quite cumbersome due to the difficulties associated with separating **2-9** and **2-5**. Additionally, amounts of trisilyl **2-8** increased on larger scale. Ultimately, it was found that compound **2-5** could be synthesized with >15:1 selectivity over compound **2-9** by a two step process (Figure 2.2b). First, compound **2-7** was treated with a small excess of tert-butyldimethylsilyl chloride (TBSCl) to afford the monosilylated **2-10**. Next, compound **2-10** could be treated with TBSOTf in the absence of base

to afford the desired compound **2-5** in 91% yield, and the undesired **2-9** in an isolated yield of 6%. This approach provided compound **2-5** on multi-gram scale and was key for the synthesis moving forward. The exact origins of the observed selectivity for the 3' silylated product are still not well understood, though under conditions involving base an isomerization likely happens based on analysis of the reaction at early time points. This isomerization is obviously diminished in the optimized procedure without added base.

2.1.4 Synthesis of Glycosyl Donor

2.1.4.1 Glycosyl Donor Synthesis with Chromatography

Upon synthesis of protected adenosine **2-5**, attention was turned to creation of an orthogonally protected ribose for chemical glycosylation. After several unsuccessful approaches, glycosyl donor **2-6** was ultimately arrived at and was determined to possess the appropriate characteristics with which to move forward (Figure 2.3). The glycosylation was found to be sensitive to steric effects, especially with respect to the protecting groups on the C-2 and C-3 alcohols. For instance, glycosyl donors such as **2-11** with bulky groups on the C-2 and C-3 alcohols were found to glycosylate poorly. Most glycosyl donors where the alcohols were protected with ethers, such as **2-12**, **2-13**, and **2-14**, glycosylated with reasonable efficiency. Compounds **2-12**, **2-13**, and **2-14** suffer from a lack of orthogonality of the protection of the C-5 position relative to the C-2 and C-3 alcohols and their further elaboration was found to be challenging.

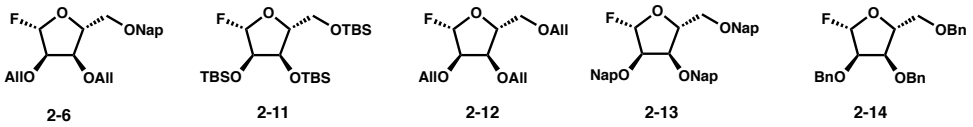


Figure 2.3. Structures of successful (**2-6**) and unsuccessful (**2-11**, **2-12**, **2-13**, and **2-14**) glycosyl donors used in these studies.

Therefore, the rationale of the final design of compound **2-6** was as follows. For the C-2 and C-3 positions, the small and non-participating allyl protecting group (All) was chosen. For protection of the C-5 alcohol, the 2-naphthylmethyl ether group (Nap) was chosen due to its ability to be removed under either mildly oxidizing or mildly reducing conditions. Protection of the C-2 and C-3 alcohols is therefore orthogonal to protection to the C-5 alcohol in this system; either can be deprotected in the presence of the other. Finally, the glycosyl fluoride was chosen as the activating group for the anomeric position due to its mild activation by fluorophilic promoters[15] and literature reports of high selectivity for producing α-furanoses[16].

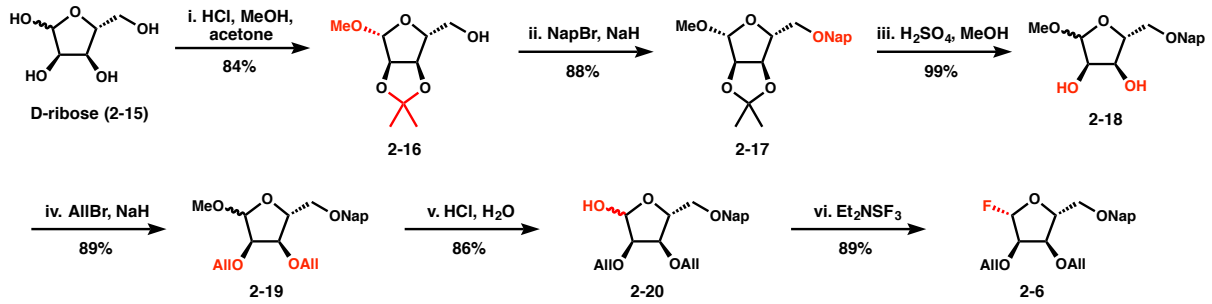


Figure 2.4. Synthesis of glycosyl donor **2-6**. i. HCl, MeOH, acetone, reflux, 4 h, 84% ii. NapBr (1 eq), NaH (1.5 eq), DMF, 0°C to rt, 36 h, 88% iii. H_2SO_4, MeOH, 100°C, 5 h, 99% iv. AllBr (4 eq), NaH (4 eq), DMF, 0°C to rt, 12 h, 89% v. HCl, dioxane, 140°C, 2 h, 86% vi. Et_2NSF_3 (1.2 eq), THF, -50°C to rt, 1 h, 89%.

In the forward sense, protected ribose **2-6** was synthesized in a high-yielding six-step sequence (Figure 2.4). First, D-ribose (**2-15**) was converted to methyl isopropylidene D-ribose (**2-16**) using a literature procedure[17], effectively protecting the C-1, C-2, and C-3 hydroxyl groups while leaving the C-5 alcohol free for further elaboration. The C-5 alcohol was subsequently naphthylated using sodium hydride and naphthyl bromide to afford compound **2-**

17. To install the desired allyl groups on the C-2 and C-3 alcohols, the acetonide of compound **2-17** was methanolyzed to form **2-18**. Compound **2-18** was efficiently bis-allylated in the next step using allyl bromide and sodium hydride to afford **2-19**. Hydrolysis of the acetal protecting C-1 of compound **2-19** was performed in aqueous dioxane to give lactol **2-20**. Finally, lactol **2-20** could be easily converted to the desired glycosyl fluoride **2-6** by treatment with diethylamino sulfurtrifluoride (DAST), giving very high selectivity for the β-fluoride.

2.1.4.2 Streamlined Synthesis without Chromatography

The route described in section 2.1.3.1 was sufficient to access gram scale quantities of glycosyl donor **2-6**, but was found to be somewhat laborious due to the number of steps. As all the steps proceeded in good yield with minimal byproducts, it was thought that perhaps a route that omitted silica column chromatography purification steps could be developed. This goal was ultimately achieved (Figure 2.5). D-ribose (**2-11**) could be converted to alcohol **2-16** as described, and alcohol **2-16** could be distilled at high temperature under vacuum to give a single β-anomer. The subsequent C-5 naphthyl group installation, isopropylidene cleavage, and C-2 and C-3 allyl group installation could be performed without purifying the intermediates by column chromatography. Chromatography was only performed after hydrolysis to lactol **2-20** and the yields of lactol **2-20** from D-ribose (**2-15**) were in some cases higher using the chromatography free procedure (see Figure 2.4).

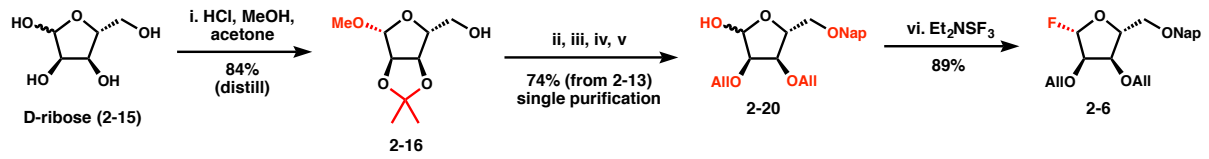


Figure 2.5. Streamlined (chromatography free) synthesis of glycosyl donor **2-6**. i. HCl, MeOH, acetone, reflux, 4 h, 84% (distill) ii. NapBr (1 eq), NaH (1.2 eq) DMF, 0°C to rt, 72 h. iii. H_2SO_4, MeOH, 110°C, 4 h. iv. AllBr (4.8 eq), NaH (5 eq), DMF, 0°C to rt, 3 h. v. aq. HCl, dioxane, reflux, 5 h, 74% (from **2-16**). vi. Et_2NSF_3 (1.2 eq), THF, -50°C to rt, 1 h, 89%.

2.1.5 Chemical Glycosylation to form Disaccharide **2-21**

Upon synthesis of compounds **2-5** and **2-6**, the formation of the desired disaccharide **2-21** via chemical glycosylation (Figure 2.6) presented the next challenge of the synthesis. As mentioned above, activation of glycosyl fluorides is frequently accomplished by use of fluorophilic promoter/co-promoter systems and a brief screen was performed to identify conditions ideal for this transformation. As a starting point, traditional systems[15] employing tin(II) chloride and silver(I) triflate or silver(I) perchlorate were examined. It was surprising to find that these conditions gave little to none of the desired product despite their use in the literature to form α-ribosyl linkages. Gratifyingly, when silver hexafluorophosphate was used in place of the other silver(I) salts, it was found that the desired disaccharide was formed in synthetically useful yields (initially about 50%). Silver hexafluorophosphate has been described as an excellent activator of thioglycosides[18], but such properties had not previously been described for glycosyl fluorides. As this procedure was scaled up, higher yields (around 70%) were obtained likely due to several factors. First, it is likely that the starting materials and reagents were rendered more anhydrous on larger (>1 g) scales than on the scales in which screening experiments were performed (>10 mg). Additionally, it was found that protecting the $AgPF_6$ from light and segregating $AgPF_6$ and $SnCl_2$ while drying the reaction components was crucial to maximize product formation.

Formation of disaccharide **2-21** and verification of synthesis of the desired steroisomer was confirmed by global deprotection to form the known compound **2-22**. Compound **2-22** was fully consistent with previously reported data[19,20]. It should be noted that in the synthesis of disaccharide **2-21**, none of the undesired β-isomer could be identified and characterized. This result is consistent with previous reports showing high α-selectivity in ribose systems with simpler glycosyl acceptors[16]. An "inside attack" model developed by Woerpel[21] and computational studies by Codee[22] suggesting that ribofuranosyl oxocarbenium ions likely adopt an E_3 conformation provide the rationale for the high stereochemical selectivity observed in our system.

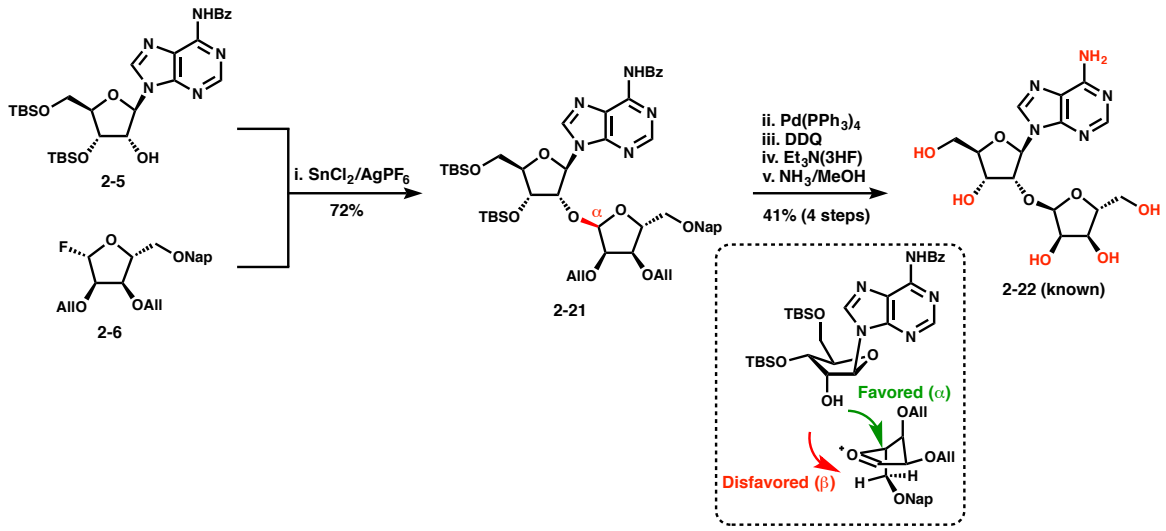


Figure 2.6. Chemical glycosylation to form disaccharide **2-21** and global deprotection to confirm its structure. i. $SnCl_2$/AgOTf (2.2/2.2 eq), CH_2Cl_2, 4 A MS, -78°C to 4 °C, 12 h, 72%. ii. $Pd(PPh_3)_4$ (20 mol%), 1,3-dimethylbarbituric acid (10 eq), MeOH, rt, 16 h. 99% iii. DDQ (1.1 eq), CH_2Cl_2/H_2O, 0°C, 5 h iv. $Et_3N{\cdot}3HF$ (5 eq), CH_2Cl_2, rt, 5 h. v. NH_3/MeOH, rt, 12 h, 42% for 3 steps. Inset: transition state rationalization for the observed α-stereochemistry.

2.1.6 Elaboration of Disaccharide **2-21** to a Suitable Coupling Piece

Upon synthesis of the key disaccharide **2-21** and verification of its structure, elaboration of this compound to form disaccharide **2-3**, the molecule possessing the requisite phosphorous functionality for fragment couplings, was undertaken (Figure 2.7). Towards this end, compound

2-21 was deallylated using $Pd(PPh_3)_4$ to afford diol **2-23**. Silylation of **2-23** using TBSOTf afforded **2-24** resulting in a net exchange of the allyl groups of **2-21** for TBS groups in **2-24**. This exchanged was performed to facilitate a global desilylation at the end of the synthesis. Tetrasilyl disaccharide **2-24** was selectivity deprotected to form a free 5' hydroxyl group by treatment with trichloroacetic acid in THF/H_2O to give alcohol **2-25**. Phosphoramidite coupling and oxidation formed phosphate **2-26**. The desired H-phosphonate **2-3** was then synthesized in a three-step process. First, the 2-naphthylmethyl group was removed from the 5'' alcohol under mildly oxidative conditions using DDQ to afford alcohol **2-27**. Treatment of alcohol **2-27** with the chlorophosphoramidite **2-28** effected formation of phosphoramidite **2-29**. Hydrolysis of phosphoramidite **2-29** gave the desired H-phosphonate **2-3** for use in the fragment couplings.

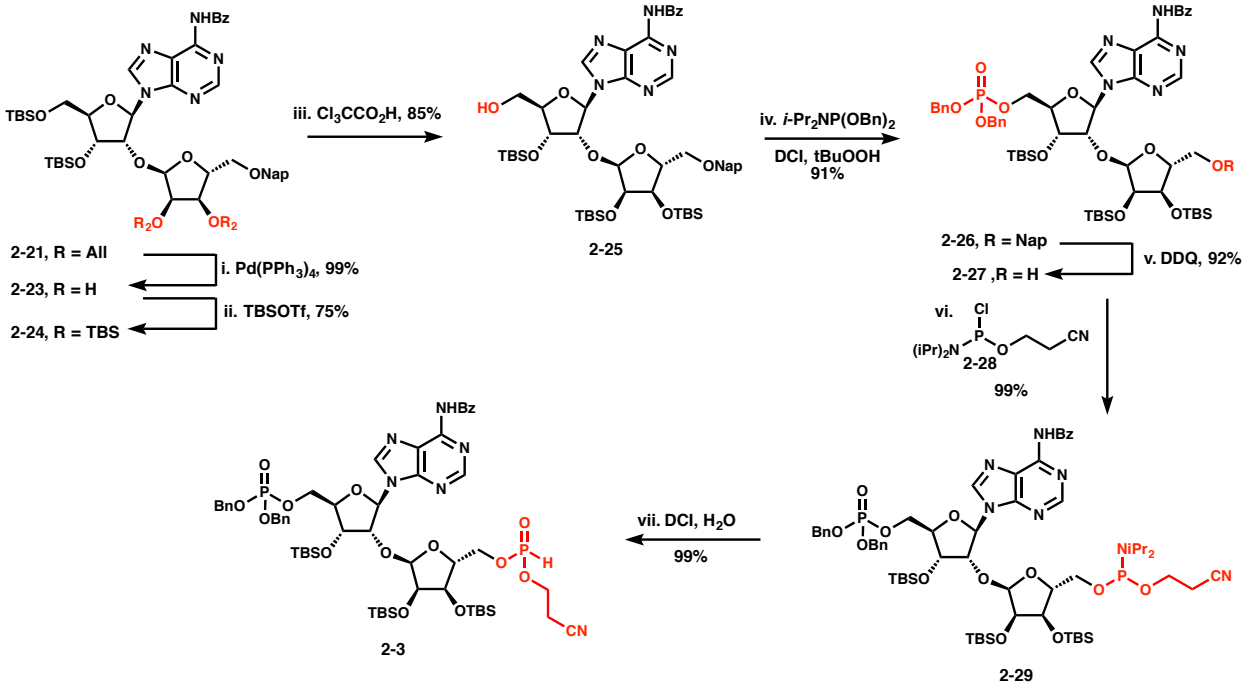


Figure 2.7. Elaboration of disaccharide **2-21** to phosphorylated disaccharide **2-3**. i. $Pd(PPh_3)_4$ (20 mol%), 1,3-dimethylbarbituric acid (10 eq), MeOH, rt, 16 h, 99%. ii. TBSOTf (3 eq), DMAP (1 eq), EtN^iPr_2 (6 eq), CH_2Cl_2, 0°C to rt, 2 h, 75%. iii. Cl_3CCO_2H (46 eq), THF/H_2O, 0°C, 5 h, 85%. iv. i-$Pr_2NP(OBn)_2$ (1.2 eq), dicyanoimidazole (1.2 eq), CH_2Cl_2/CH_3CN, 2 h, rt, then: tBuOOH (5 eq), 0°C to rt, 3 h, 91%. v. DDQ (1.5 eq), CH_2Cl_2/H_2O, 0°C, 6 h, 92%. vi. **2-28** (1.2 eq), EtN^iPr_2 (1.2 eq), THF, -78°C to rt, 3 h, 99%. vii. dicyanoimidazole (1.2 eq), H_2O (3 eq), CH_3CN, rt, 2 h, 99%.

2.1.7 Synthesis of Protected Adenosine Monophosphate **2-4**

With the challenging disaccharide **2-3** in hand, attention was turned to the synthesis of protected adenosine monophosphate **2-4** and was accomplished in a five-step sequence as follows (Figure 2.8). First, N-benzoyl adenosine (**2-7**) was trisilylated with TBSCl and imidazole via a known procedure[23] to form **2-8**. Deprotection of the primary silyl group had been previously reported and was accomplished using trichloroacetic acid in THF/water to afford alcohol **2-30**[24]. Phosphoramidite coupling of alcohol **2-30** and phosphoramidite **2-31** afforded biscyanoethyl phosphate **2-32**. Deprotection to form the initiator phosphate **2-4** was accomplished with DBU and TMSCl, with TMSCl being required to effect cleavage of both cyanoethyl groups.

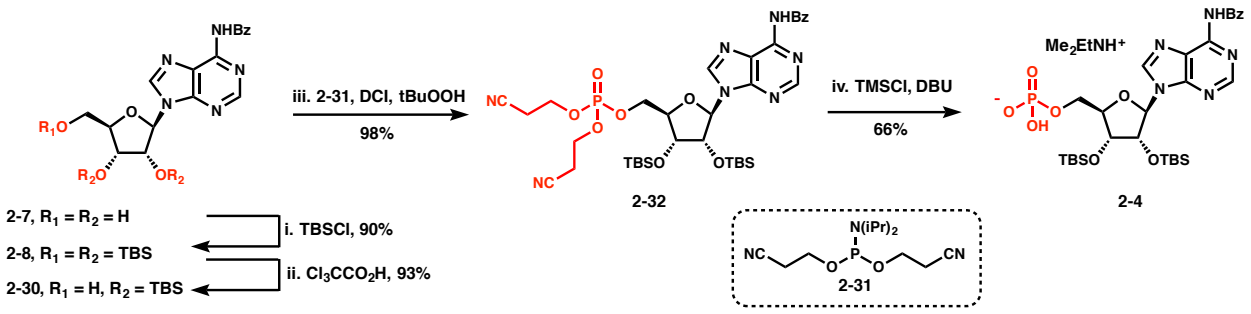


Figure 2.8. Synthesis of adenosine monophosphate **2-4**. i. TBSCl (5 eq), imidazole (5 eq), DMF, 0°C to rt, 1 h, 90%. ii. Cl_3CCO_2H (46 eq), THF/H_2O, 0°C, 3 h, 93%. iii. **2-31** (1.2 eq), dicyanoimidazole (1.2 eq), CH_2Cl_2/CH_3CN, 1.5 h, rt, then: tBuOOH (5 eq), 0°C to rt, 3 h, 98%. iv. DBU (10 eq), TMSCl (6 eq), CH_3CN, 0°C to rt, 12 h, reverse phase chromatography, ion exchange, 66%.

2.1.8 Synthesis of Protected ribose Phosphate **2-2**

In contrast to the initiator phosphate **2-4**, accessing a terminating ribose phosphate proved to be more challenging. Several protecting group schemes were attempted (Figure 2.9) and the combination shown for compound **2-2** was ultimately successful. Use of **2-33** led to low

coupling yields, while the benzyl groups on the secondary alcohols of **2-34** were hard to remove by hydrogenolysis. The TBS group was found to be too labile for protection of the anomeric position as in **2-35**. Ultimately, compound **2-2** was determined to be ideal. The choice of TBS groups for protection of the C-2 and C-3 alcohols was clear due to their presence on the other alcohols of the molecule. The tert-butyldiphenylsilyl (TBDPS) group was chosen for the anomeric position due to its increased acid stability relative to the TBS group, and the propensity for anomeric substituents to be labile under acidic conditions as was seen for **2-35**.

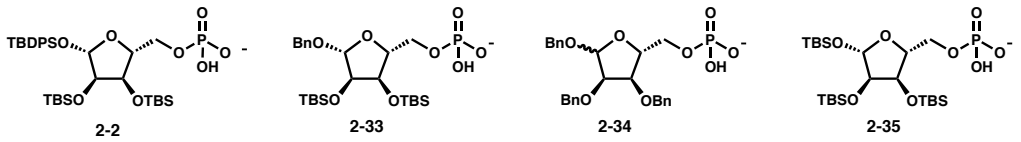


Figure 2.9. Structures of ribose phosphates used for this route. **2-2** was ultimately successful, while **2-33**, **2-34**, and **2-35** failed for various reasons.

The route to synthesize compound **2-2** (Figure 2.10) commenced from commercially available D-ribonolactone (**2-36**). The C-5 alcohol was protected with a trityl group using a known procedure,[25] followed by silylation with TBSOTF to afford the known lactone **2-37**[26]. Reduction using *i*Bu_2AlH gave lactol **2-38**, and the TBDPS group was installed using TBDPSCl and imidazole in DMF. Deprotection of **2-39** to afford alcohol **2-40** was accomplished using using trifluoroacetic acid with triethylsilane as a trityl scavenger. Phosphoramidite coupling of alcohol **2-40** with dibenzyl diisopropylphosphoramidite afford protected phosphate **2-41** and hydrogenolysis gave the desired ribose terminator **2-2** in good yield.

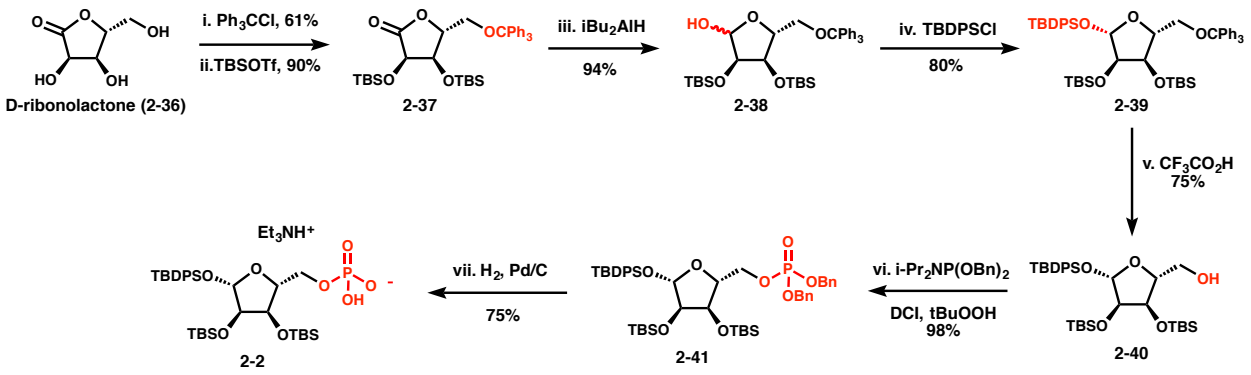


Figure 2.10. Synthesis of ribose phosphate **2-2**. i. Ph_3CCl (1.2 eq), DMAP (0.2 eq), pyridine, 70°C, 16 h, 61%. ii. TBSOTf (3 eq), 2,6-lutidine (6 eq), CH_2Cl_2, 0°C, 1.5 h, 90%. iii. $^{i}Bu_2AlH$ (1.3 eq), CH_2Cl_2, -78°C, 1 h, 94%. iv. TBDPSCl (2 eq), imidazole (2.5 eq), DMF, 70°C, 18 h, 80%. v. CF_3CO_2H (2 eq), Et_3SiH (4 eq), CH_2Cl_2, 0°C, 1 h, 75%. vi. i-$Pr_2NP(OBn)_2$ (1.4 eq), dicyanoimidazole (1.4 eq), CH_3CN, 1.5 h, rt, then: tBuOOH (5 eq), rt, 1 h, 98%. vii. H_2, Pd/C, Et_3N (10 eq), tBuOH/H_2O, rt, 16 h, 75%.

2.1.9 Fragment Couplings and Completion of the ADP-ribose Dimer

With the requisite coupling partners in hand, the fragment coupling reactions (Figure 2.11) were performed. Treatment of H-phosphonate **2-3** with N-chlorosuccinimide in the presence of phosphate **2-4** and Hunig's base followed by DBU afforded pyrophosphate **2-42**. This reaction likely proceeds through an oxidative chlorination to form a highly electrophilic chlorophosphate, followed by condensation with the phosphate to form a pyrophosphate triester. Deprotection of the pyrophosphate triester with DBU gives the desired pyrophosphate diester. We found this method to be a rapid and high-yielding method for pyrophosphate formation, and believe that it is superior to most methods for pyrophosphate synthesis in the literature. This method has been broadly used in our lab for the synthesis of pyrophosphates in molecules such as ADP-HPD and also ADP-ribose-based probes.

With the first pyrophosphate of dimeric ADP-ribose formed, the benzyl phosphate **2-42** was deprotected by hydrogenolysis to give deprotected phosphate **2-43**. Attempts to couple phosphate **2-43** with a terminating H-phosphonate under the oxidative chlorination conditions

described above were unsuccessful. Subsequently, similar methods that rely on highly electrophilic phosphorous species for coupling in the presence of another pyrophosphate have also been reported to be unsuccessful in the solution phase[27]. As a result, a more traditional phosphorimidazolide coupling was used. In this procedure, phosphate **2-2** was activated as a phosphorimidazolide using carbonyl diimidazole and coupled with phosphate **2-43** in DMF using zinc(II) chloride as a promoter. Removal of the zinc salts using EDTA and reverse phase purification gave protected ADP-ribose dimer **2-44** in acceptable yield. Deprotection of compound **2-44** with methanolic ammonia (to remove the benzoyl groups) and tetrabutylammonium fluoride (to remove the silyl protecting groups) gave the ADP-ribose dimer (**2-1**) in good yield after C_{18} ion-pairing purification. Extensive 2D NMR spectroscopic analysis was performed to verify the structure of dimer **2-1**. Further structural confirmation was provided by the ability of the compound to be processed by PARG as well as its co-crystal structure with mutant PARG (discussed in chapter 3).

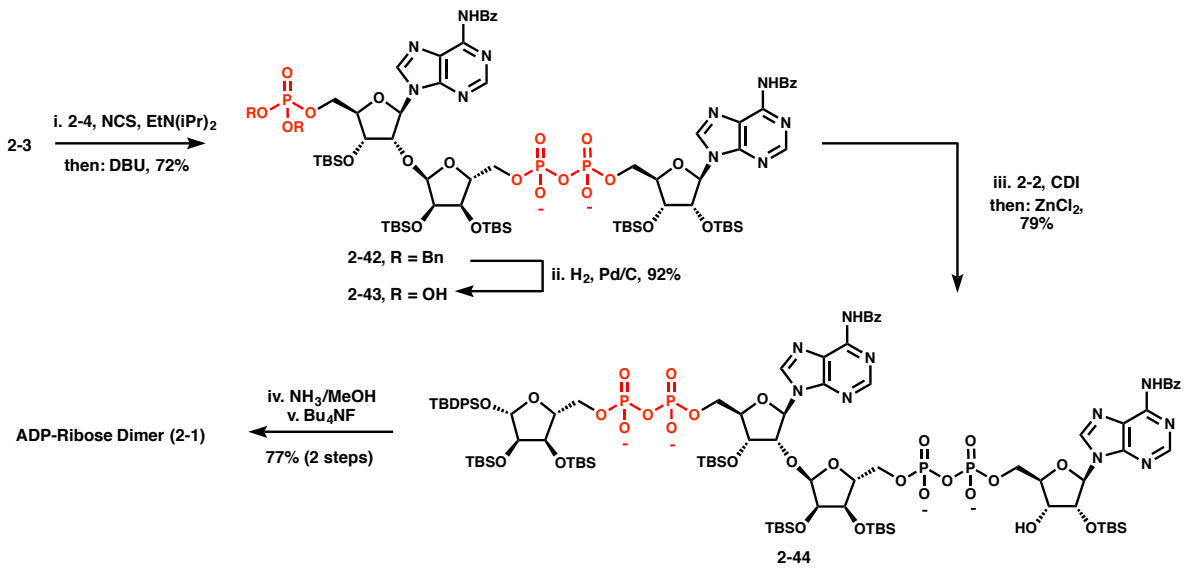


Figure 2.11. Fragment couplings and completion of the ADP-ribose dimer (**2-1**). i. NCS (4 eq), EtN^iPr_2 (4 eq), CH_3CN, rt, 20 min, then: DBU (10 eq), CH_3CN, rt, 20 min, 72%. ii. H_2, Pd/C, Et_3N (10 eq), tBuOH/H_2O, rt, 16 h, 92%. iii. **2-2**, carbonyldiimidazole (10 eq), Et_3N (4 eq), pyridine, rt, 2 h, then: **2-43**, $ZnCl_2$ (8 eq), DMF, rt, 96 h, reverse phase purification, ion exchange, 79%. iv. NH_3/MeOH. v. Bu_4NF (33 eq), THF, rt, 3 h, ion pairing chromatography, ion exchange, 77%.

2.2 Synthesis of the Propargyl ADP-ribose Dimer

2.2.1 Motivation and Retrosynthesis

The opportunity to synthesize derivatives of compound **2-1** presented by the convergent nature of the synthesis was immediately recognized as a method through which biochemical tools could be accessed. Towards this end, it was desired to synthesize an alkyne-containing ADP-ribose dimer (Figure 2.12, **2-45**) that could be used with Huisgen 1,3-dipolar cycloaddition chemistry to append fluorophores and other useful molecules to the ADP-ribose dimer. It was envisioned that this ADP-ribose dimer derivative would originate from a pyrophosphate coupling between the previously described phosphate **2-43** and a new alkyne-containing ribose phosphate **2-46** that could be synthesized via propargylation of lactol **2-38**.

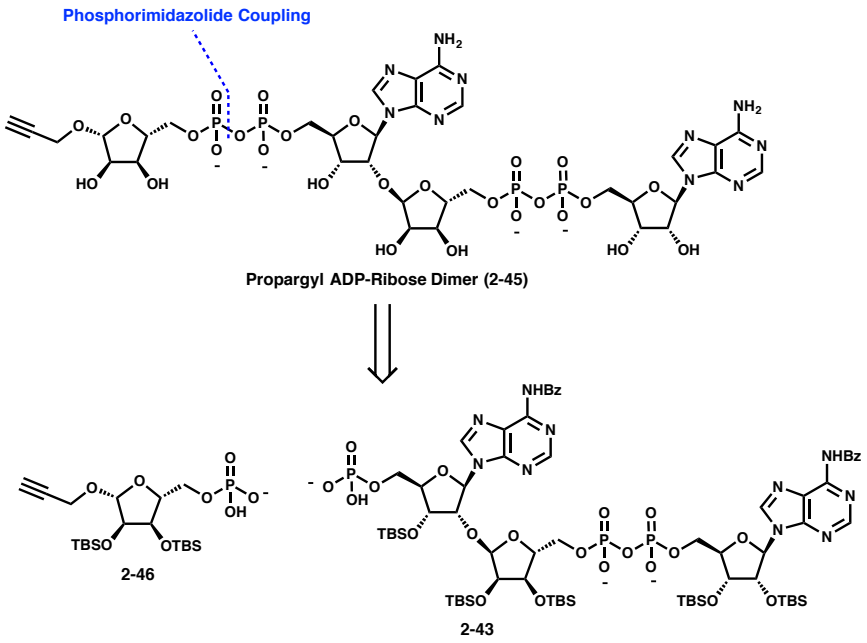


Figure 2.12. Retrosynthetic analysis of the propargyl ADP-ribose dimer (**2-45**).

2.2.2 Synthesis of Propargyl Ribose Phosphate

Synthesis of propargyl ribose phosphate **2-46** (Figure 2.13) was relatively straightforward and involved lactol **2-38**, an intermediate on the synthesis to ribose terminator **2-2**. Propargylation of lactol **2-38** using propargyl bromide and sodium hydride gave compound **2-47** as a single β-anomer, likely as a result of the steric bulk of the TBS groups. Deprotection of the trityl group of **2-47** was performed using trifluoroacetic acid to give alcohol **2-48**. In contrast to the synthesis of protected phosphate **2-41** on route to the ADP-ribose dimer, alcohol **2-48** was coupled with fluorenylmethyl phosphoramidite **2-49** rather than the analogous dibenzyl reagent due to the susceptibility of the propargyl group to reduction via hydrogenation. Deprotection of **2-50** with triethylamine gave the desired phosphate **2-46**.

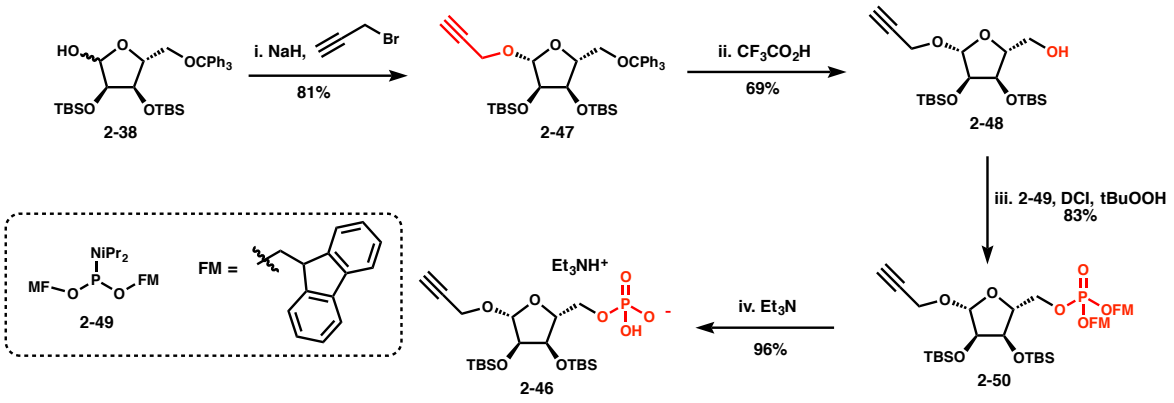


Figure 2.13. Synthesis of propargyl ribose phosphate **2-46**. i. NaH (2 eq), propargyl bromide (1.2 eq), DMF, 0°C, 1 h, 81%. ii. CF_3CO_2H (2 eq), Et_3SiH (4 eq), CH_2Cl_2, 0°C, 40 min, 69%. iii. **2-49** (1.5 eq), dicyanoimidazole (1.5 eq), CH_3CN/CH_2Cl_2, 1.5 h, rt, then: tBuOOH (5 eq), rt, 1 h, 83%. iv. Et_3N/CH_3CN, rt, 14 h, 96%.

2.2.3 Fragment Couplings and Completion of the Propargyl ADP-ribose Dimer (**2-45**)

Upon synthesis of propargyl ribose terminator **2-46**, assembly of the fragments to produce the propargyl ADP-ribose dimer (**2-45**) was undertaken (Figure 2.14). A phosphorimidazolide coupling between phosphates **2-43** and **2-46** was performed to give fully protected propargyl ADP-ribose dimer **2-51**. Global deprotection gave propargyl ADP-ribose dimer **2-45**.

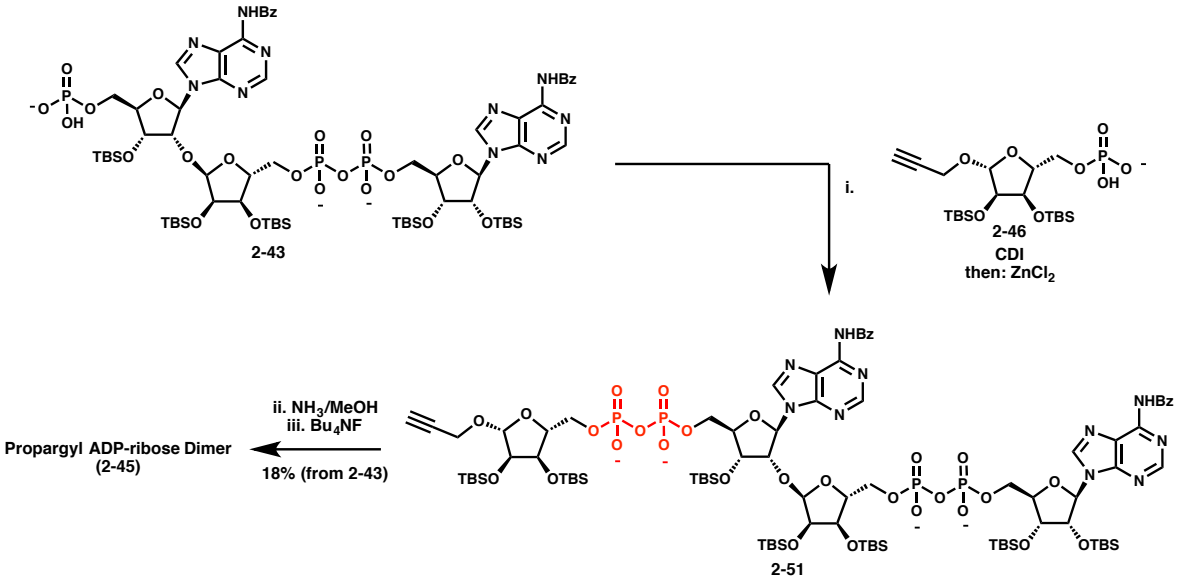


Figure 2.14. Fragment couplings to access propargyl ADP-ribose dimer **2-45**. i. **2-46**, carbonyldiimidaozle (5 eq), Et_3N (14 eq), CH_3CN, rt, 2 h, then: **2-46**, $ZnCl_2$ (8 eq), rt, 30 h. ii. NH_3/MeOH, rt, 14 h. iii. Bu_4NF, THF, rt, 3 h, ion pairing chromatography 18% from **2-43**.

2.2.4 Synthesis of ADP-ribose Dimer Tool Molecules

Upon successful synthesis of the propargyl ADP-ribose dimer (**2-45**), its use in the synthesis of tool molecules using copper-catalyzed Husigen 1,3-dipolar cycloaddition chemistry was demonstrated. Biotinylated (**2-52**) and Cy3-containing (**2-53**) ADP-ribose dimers were synthesized using this chemistry (Figure 2.15).

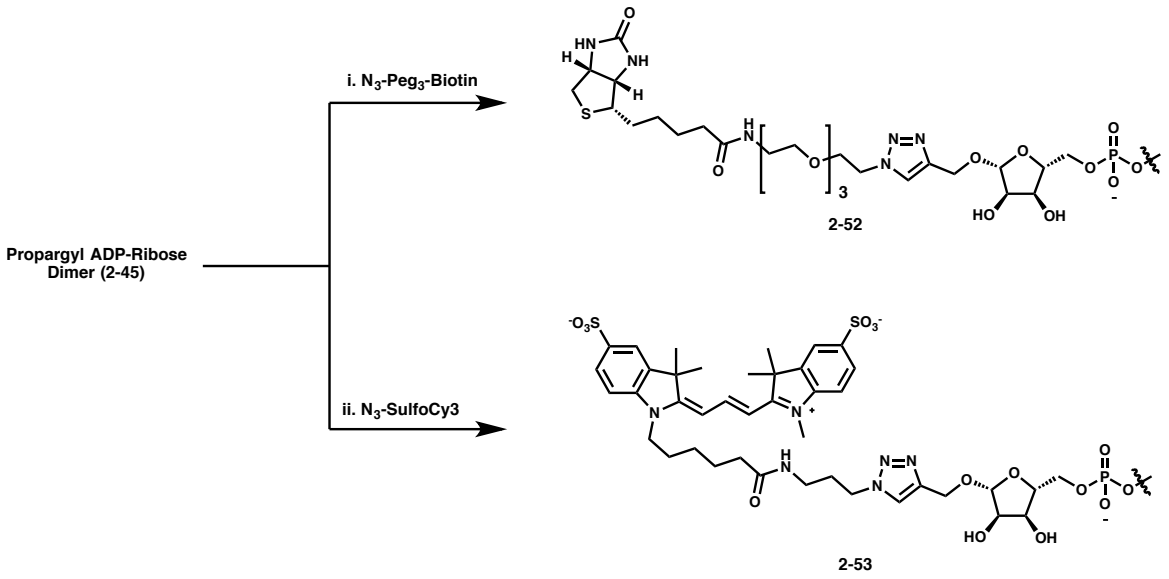


Figure 2.15. Synthesis of biotinylated (**2-52**) and Cy3-ADP-ribose (**2-53**) dimers. i. Biotin-Peg_3-N_3 (2 eq), $CuSO_4{\cdot}5H_2O$ (2 eq), Cu wire, H_2O, rt, 16 h. ii. SulfoCy3-N_3, $CuSO_4{\cdot}5H_2O$ (0.6 eq), Cu wire, H_2O, rt, 16 h.

2.3 Summary

The synthesis of the ADP-ribose dimer (**2-1**) and the propargyl ADP-ribose dimer (**2-38**) were achieved. Key to the route to both compounds was the α-selective glycosylation to form the disaccharide repeating unit **2-5**. Additionally, the fast and high yielding pyrophosphate coupling to form the first pyrophosphate en route to the ADP-ribose dimers substantially improved material throughput with respect to more traditional methods of pyrophosphate formation. While this synthetic route provides access to dimers **2-1** and **2-45**, its major shortcoming is the low yields associated with the imidazolide coupling reactions used for the formation of the second pyrophosphate (synthesis of **2-44** from **2-43** or **2-51** from **2-43**) and the inability of the H-phosphonate coupling method (used successfully for the synthesis of **2-42** from **2-3**) to proceed in the presence of another pyrophosphate in the molecule. Therefore, while this route is successful in providing dimers **2-1** and **2-45** and such molecules enable important biological

experiments (chapter 3), alternative approaches (such as those described in chapter 4) are necessary for the synthesis of longer oligomers.

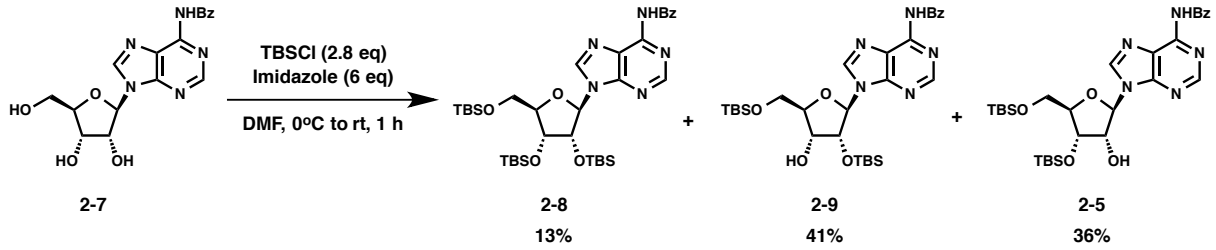
2.4 Experimental

All reactions were run in flame or oven dried glassware under an atmosphere of dry nitrogen unless otherwise noted. Acetonitrile, tetrahydrofuran, methanol and methylene chloride used in reactions were obtained from a solvent dispensing system. Diethyl ether was distilled from sodium metal. 4 Å molecular sieves were dried at 200 ^{o}C on high vacuum overnight. Pyridine and Diazabicyclo[5.4.0]undec-7-ene were distilled from CaH_2 and stored on 4 Å molecular sieves. 1,2-Dichloroethane and d_3-acetonitrile were dried over 4 Å molecular sieves. All other reagents were of standard commercial purity and were used as received. Tetrakis(triphenylphosphine)palladium(0) was prepared using the procedure of Malpass, *et al.*[28] and was packaged in a glove-bag under nitrogen.

Analytical thin-layer chromatography was performed on EMD Merck silica gel plates with F254 indicator. Plates were visualized with UV light (254 nm) or staining with p-anisaldehyde. Silica gel for column chromatography was purchased from Sorbent Technologies (40-75 mm particle size).

Unless otherwise indicated, ^{1}H, ^{13}C, ^{19}F, and ^{31}P NMR spectra were recorded at 500, 125, 470 and 203 MHz, respectively. ^{1}H and ^{13}C NMR spectra were referenced to tetramethylsilane or the residual solvent peak as reported by Fulmer *et al*.[29] ^{13}C Spectra taken in D_2O contain 0.05% EtOH (17.47, 58.05 ppm) as an internal standard for referencing. ^{19}F NMR spectra were referenced using C_6F_6 as an internal standard (-164.9 ppm). ^{31}P NMR spectra were externally referenced to 85% H_3PO_4 (0.00 ppm) in water. Chemical shifts are reported in ppm and multiplicities are reported as s (singlet), d (doublet), t (triplet), q (quartet), p (pentet), h (hextet),

hep (heptet), m (multiplet), and br (broad). Mass spectrometry analysis was performed by the University of Illinois Mass Spectrometry Center or by direct injection on an Agilent 6230 LC/MS TOF for samples run in negative ion mode. The LC/MS assay was performed on the Agilent 6230 LC/MS TOF system with a 1.8 mm, 2.1x50 mm Agilent ZORBAX Eclipse Plus C18 column and is described in more detail below. Analytical HPLC analysis was performed on a Waters e2695 separations module with a Waters 2489 UV detector using a 5 mm, 4.6x150 mm Waters XBridge BEH130 HPLC column and is described in more detail below. Other preparative C18 chromatography was performed using a Teledyne Isco CombiFlash Rf system with CombiFlash Gold columns using a gradient of H_2O/CH_3CN beginning with 95% H_2O:5% CH_3CN, ramping to 65% H_2O:35% CH_3CN over 6 min, ramping to 100% CH_3CN for 3 min, and holding 100% CH_3CN for 5 min. Optical rotations were measured using a Jasco DIP-360 digital polarimeter in either EtOH or $CHCl_3$ with concentrations reported in g/dL. Infrared spectra were recorded on a Mattson Galaxy 5020 spectrophotometer with NaCl cells. Peaks are reported in cm^{-1}.



2-7 | 2-8 13% | 2-9 41% | 2-5 36%

Unselective synthesis of compounds 2-8, 2-9 and 2-5: A modified version of the procedure of Cole *et al.*[13] was followed. To a 250 mL round bottom flask was added N-Benzoyladenosine (**2-7**) (10.0 g, 26.9 mmol, 1 eq). Compound **2-7** was evaporated three times with dry CH_3CN on a nitrogen-filled rotary evaporator and dried on high vacuum over P_2O_5. The starting material was dissolved in DMF (25 mL) and cooled to 0°C. Imidazole (11.0 g, 161.4 mmol, 6 eq) was added, followed by a solution of TBSCl (11.4 g, 75.4 mmol, 2.8 eq) in DMF (25 mL) at 0 °C over 30

min. The reaction was allowed to warm to room temperature and stirred for 30 min. The reaction was poured into saturated aqueous NH_4Cl, extracted three times EtOAc, filtered through Na_2SO_4 and concentrated. The products were purified by silica column chromatography to isolate compounds **2-8** (2.47 g, 13%), **2-9** (6.60 g, 41%) and **2-5** (5.89 g, 36%).

Compound **2-8**

^{1}H NMR (500 MHz, $CDCl_3$) δ 9.50 (s, 1H, NHBz), 8.82 (s, 1H, H-8), 8.35 (s, 1H, H-2), 8.05 (d, J = 7.5 Hz, 2H, Bz), 7.59 (t, J = 7.4 Hz, 1H, Bz), 7.51 (dd, *J* =7.6 Hz, 2H, Bz), 6.12 (d, *J* = 5.3 Hz, 1H, H-1'), 4.68 (t, J = 4.8 Hz, 1H, H-2'), 4.30 (appt, J = 3.9 Hz, 1H, H-3'), 4.15 (q, *J* = 3.4 Hz, 1H, H-4'), 4.03 (dd, *J* = 11.4, 3.9 Hz, 1H, H-5'a), 3.80 (dd, *J* = 11.4, 2.8 Hz, 1H, H-5'b), 0.96 (s, 9H, -tBu), 0.93 (s, 9H, -tBu), 0.78 (s, 9H, -tBu), 0.15 (s, 3H, -CH_3), 0.14 (s, 3H, -CH_3), 0.10 (s, 3H, -CH_3), 0.05 (s, 3H, -CH_3), -0.05 (s, 3H, -CH_3), -0.27 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 164.9, 152.9, 151.7, 149.8, 141.8, 134.1, 132.8, 128.9, 128.1, 123.0, 88.5, 86.0, 76.1, 72.2, 62.7, 26.2, 26.0, 25.8, 18.7, 18.2, 18.0, -3.5, -4.3, -4.5, -4.6, -4.9, -5.2.

HRMS (ESI) m/*z* calcd for $C_{35}H_{60}N_5O_5Si_3$ ($[M+H]^+$) 714.3902, found 714.3909.

IR (neat) ν 2929, 2857, 1698, 1611, 1580, 1454, 1254, 1072, 837, 776 cm^{-1}.

$[a]_D^{23}$ -35.9 (*c* = 1.23, $CHCl_3$)

Compound 2-9 (R_f 0.30 in 50:50 hexanes:ethyl acetate)

^{1}H NMR (500 MHz, $CDCl_3$) δ 9.10 (s, 1H, NHBz), 8.80 (s, 1H, H-8), 8.41 (s, 1H, H-2), 8.01 (d, *J* = 7.7 Hz, 2H, Bz), 7.62 – 7.55 (m, 1H, Bz), 7.51 (t, *J* = 7.9 Hz, 2H, Bz), 6.18 (d, *J* = 5.0 Hz, 1H, H-1'), 4.64 (appt, *J* = 5.0 Hz, 1H, H-2'), 4.28 (appq, *J* = 4.1 Hz, 1H, H-3'), 4.22 (appq, *J* = 2.5 Hz, 1H, H-4'), 4.01 (dd, *J* = 11.5, 2.5 Hz, 1H, H-5'a), 3.86 (dd, *J* = 11.5, 2.4 Hz, 1H, H-5'b),

2.79 (d, *J* = 4.2 Hz, 1H, -OH), 0.94 (s, 9H, -tBu), 0.83 (s, 9H, -tBu), 0.14 (s, 3H, -CH_3), 0.13 (s, 3H, -CH_3), -0.06 (s, 3H, -CH_3), -0.15 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 164.7, 153.0, 151.8, 149.6, 141.4, 133.9, 132.8, 129.0, 127.9, 123.1, 88.3, 85.5, 77.2, 71.4, 63.2, 26.2, 25.7, 18.6, 18.0, -4.9, -5.1, -5.2, -5.3.

HRMS (ESI) m/*z* calcd for $C_{29}H_{46}N_5O_5Si_2$ ($[M+H]^+$) 600.3038, found 600.3039

IR (neat) ν 3365, 2947, 2926, 2850, 1701, 1611, 1580, 1455, 1247, 1136, 1063, 837, 782, 706 cm^{-1}.

$[a]_D^{23}$ -35.0 (*c* = 0.88, $CHCl_3$)

Compound 2-5 (R_f 0.16 in 50:50 hexanes:ethyl acetate)

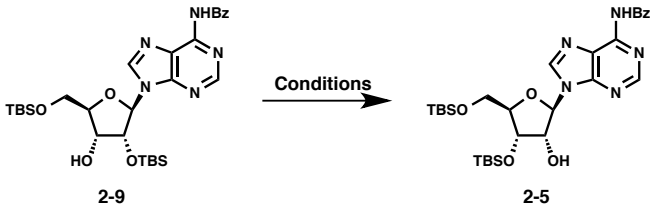
^{1}H NMR (500 MHz, $CDCl_3$) δ 9.08 (s, 1H, NHBz), 8.78 (s, 1H, H-8), 8.27 (s, 1H, H-2), 8.01 (d, *J* = 7.3 Hz, 2H, Bz), 7.59 (t, *J* = 7.4 Hz, 1H, Bz), 7.50 (t, *J* = 7.7 Hz, 2H, Bz), 6.09 (d, *J* = 4.7 Hz, 1H, H-1'), 4.63 – 4.54 (m, 2H, H-2'+H-3'), 4.13 (appq, *J* = 3.3 Hz, 1H, H-4'), 3.93 (dd, *J* = 11.4, 3.5 Hz, 1H, H-5'a), 3.78 (dd, *J* = 11.4, 2.9 Hz, 1H, H-5'b), 3.21 (d, *J* = 6.6 Hz, 1H, -OH), 0.95 (s, 9H, -tBu), 0.89 (s, 9H, -tBu), 0.17 (s, 6H, -CH_3), 0.08 (s, 3H, -CH_3), 0.05 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 164.7, 152.8, 151.8, 149.6, 141.7, 133.9, 132.9, 129.0, 128.0, 123.5, 89.1, 85.7, 75.3, 71.8, 62.5, 26.1, 25.9, 18.5, 18.2, -4.5, -4.7, -5.3, -5.4.

HRMS (ESI) m/*z* calcd for $C_{29}H_{46}N_5O_5Si_2$ ($[M+H]^+$) 600.3038, found 600.3032

IR (neat) ν 3310, 2928, 2864, 1704, 1613, 1580, 1461, 1253, 1073, 836, 775, 702 cm^{-1}.

$[a]_D^{23}$ -20.5 (*c* = 1.13, $CHCl_3$)

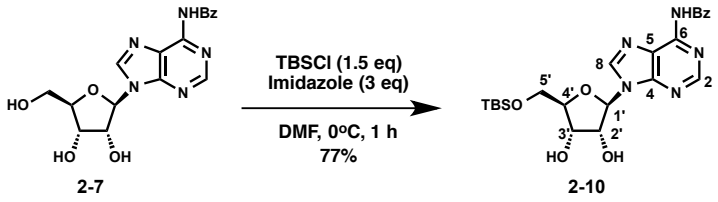


Isomerization of compound 2-9 to 2-5

To a 7 mL reaction vial was added compound **2-9** (100 mg, 0.17 mmol, 1 eq) and MeOH or EtOH (1 mL). The solution was stirred at 30 ^{o}C for 48 h, concentrated and purified by silica column chromatography to yield compound **2-5** (45 mg, 45% for MeOH treatment or 40 mg, 40% for EtOH treatment).

Alternatively, compound **2-9** (100 mg, 0.17 mmol, 1 eq) was added to a 7 mL scintillation vial, followed by imidazole (22.8 mg, 0.33 mmol, 2 eq) and DMF (2 mL). The reaction was stirred at 30 ^{o}C for 48 h, at which point a saturated aqueous NH_4Cl solution was added. The reaction was extracted three times with EtOAc, dried through Na_2SO_4, concentrated, and purified by silica column chromatography to yield compound **2-5** (50.4 mg, 50%).

Compound **2-9** could be recovered in nearly equal quantities as the desired compound (**2-5**) in all of the above treatments.



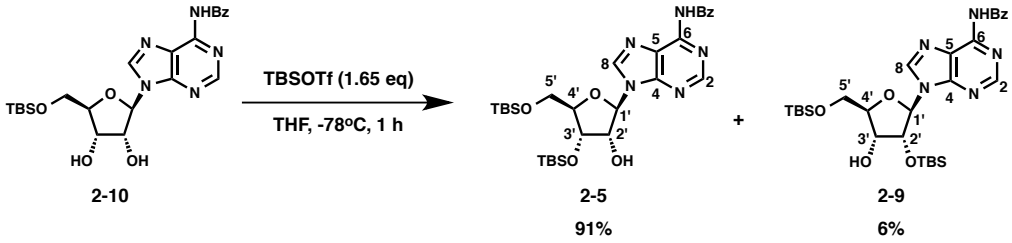
Compound 2-10: To a 0 ^{o}C solution of *N*-Benzoyladenosine (**2-7**) (10.0 g, 26.9 mmol, 1 eq) in DMF (150 mL) was added imidazole (5.5 g, 80.8 mmol, 3 eq). Next, TBSCl (6.09 g, 40.4 mmol, 1.5 eq) was added in DMF (20 mL) and the reaction was stirred at 0^{o}C for 1 h. The reaction was poured into saturated aqueous NH_4Cl, extracted three times with EtOAc, washed twice with

saturated aqueous NH_4Cl, dried through Na_2SO_4, and concentrated. Water was added to precipitate the product and the solid was washed twice with water. The solid was collected; EtOH was added and evaporated to remove residual water. The product was purified by silica column chromatography (R_f=0.31 in 25:75 Hexanes:EtOAc) to isolate compound **2-10** as a white foam (10.1 g, 77%). Compounds **2-5** (822 mg, 5%) and **2-9** (872 mg, 5%) were also isolated as minor products. Detailed characterization data for **2-5** and **2-9** is provided below.

^{1}H NMR (500 MHz, CD_3OD) δ 8.72 (s, 1H, H-8), 8.67 (s, 1H, H-2), 8.09 (d, *J* = 7.4 Hz, 2H, Bz), 7.66 (t, *J* = 7.4 Hz, 1H, Bz), 7.57 (t, *J* = 7.7 Hz, 2H, Bz), 6.19 (d, J = 4.3 Hz, 1H, H-1'), 4.66 (dd, *J* = 4.9, 4.2 Hz, 1H, H-2'), 4.41 (dd, *J* = 5.2, 4.9 Hz, 1H, H-3'), 4.16 (ddd, *J* = 5.2, 3.1, 3.1 Hz, 1H, H-4'), 4.04 (dd, *J* = 11.6, 3.1 Hz, 1H, H-5'a), 3.91 (dd, *J* = 11.6, 3.1 Hz, 1H, H-5'b), 0.94 (s, 9H, -tBu), 0.13 (s, 6H, $-CH_3$).

^{13}C NMR (125 MHz, CD_3OD) δ 167.9, 153.2, 153.0, 151.0, 143.8, 134.9, 133.8, 129.7, 129.4, 124.9, 90.3, 86.4, 76.3, 71.2, 63.8, 26.5, 19.3, -5.3, -5.3.

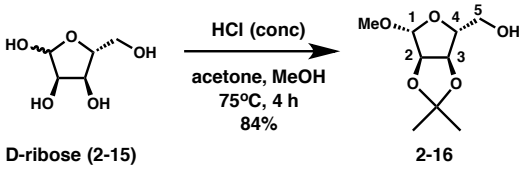
HRMS (ESI) m/*z* calcd for $C_{23}H_{32}N_5O_5Si$ ($[M+H]^+$) 486.2173, found 486.2173.



2-10 | 2-5 91% | 2-9 6%

Selective synthesis of compound 2-5: To a 250 mL round bottomed flask containing compound **2-10** (1.70 g, 3.5 mmol, 1 eq) was added THF (35 mL), and the solution was cooled to -78°C. Next, TBSOTf (0.88 mL, 3.85 mmol, 1.1 eq) was added and the solution was stirred at -78°C for 30 min. After 30 min, additional TBSOTf (0.44 mL, 1.92 mmol, 0.55 eq) was added and stirred for 30 min at -78°C. The reaction was quenched at -78°C with saturated aqueous $NaHCO_3$ and

slowly warmed to room temperature. The aqueous layer was extracted once with ethyl acetate, once with methylene chloride and once with chloroform. The combined organic layers were dried though Na_2SO_4, concentrated and purified by silica column chromatography to give known[13] compounds **2-5** (1.92 g, 91%) **2-9** as a minor byproduct (0.12 g, 6%) as white foams.



Compound 2-16: A modified version of the procedure of Paquette *et al*.[17] was followed. Specifically, to a 250 mL round bottom flask was added D-ribose (**2-15**) (10.0 g, 66.1 mmol, 1 eq), methanol (40 mL), and acetone (40 mL). Next, concentrated HCl (1 mL) was added at room temperature, the flask was sealed, and the reaction was heated to 75 ^oC for 4 h. The solution was cooled to room temperature and quenched with pyridine (10 mL). The solution was concentrated, diluted with EtOAc, and poured into a saturated aqueous $NaHCO_3$ solution. The aqueous layer was extracted ten times with EtOAc, the combined organic layers were dried over Na_2SO_4, concentrated, and purified by distillation under high vacuum (bp 110 ^oC at 1 torr, R_f 0.34, 50:50 hexanes:ethyl acetate). Compound **2-16** was obtained as a colorless oil (11.3 g, 84%).

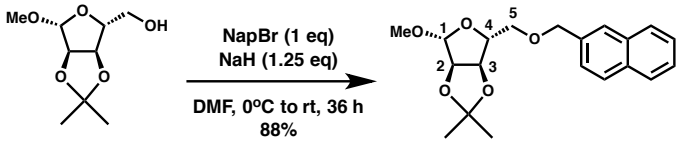
^{1}H NMR (499 MHz, $CDCl_3$) δ 4.97 (s, 1H, H-1), 4.84 (d, *J* = 5.9 Hz, 1H, H-2), 4.59 (d, *J* = 5.9 Hz, 1H, H-3), 4.43 (bs, 1H, H-4), 3.70 (ddd, *J* = 12.6, 2.4, 2.7 Hz, 1H, H-5a), 3.65-3.57 (m, 1H, H-5b), 3.44 (s, 3H, -OCH_3), 3.21 (dd, *J* = 10.7, 2.7 Hz, 1H,-OH), 1.49 (s, 3H, -CH_3), 1.32 (s, 3H, -CH_3).

^{13}C NMR (126 MHz, $CDCl_3$) δ 112.3, 110.3, 88.6, 86.0, 81.7, 64.2, 55.6, 26.6, 25.0.

HRMS (ESI) m/*z* calcd for $C_9H_{16}O_5Na$ ($[M+Na]^+$) 227.0895, found 227.0896.

IR (neat) ν 3457, 2986, 2934, 2836, 1455, 1380, 1276, 1208, 1162, 1097, 1045, 1009, 963, 869 cm^{-1}.

$[a]_D^{21}$ -73.5 (*c* = 2.71, $CHCl_3$)



2-16 2-17

Compound 2-17: To a 500 mL round bottom flask was added DMF (180 mL) and compound **2-16** (10.0 g, 49.0 mmol, 1.1 eq). The reaction was cooled to 0 ^oC. Next, NaH (1.34 g, 55.6 mmol, 1.25 eq) was added in one portion and stirred for 15 min at 0 ^oC. Lastly, 2-bromomethyl-naphthalene (9.84 g, 44.5 mmol, 1 eq) was added portionwise and slowly warmed to room temperature. The reaction was stirred at room temperature for 36 h. The reaction was then cooled to 0 ^oC, quenched with saturated aqueous NH_4Cl, extracted with EtOAc twice, the organic layer was dried over Na_2SO_4, concentrated and purified by silica gel chromatography. Compound **2-17** was obtained as a white solid (13.54 g, 88%).

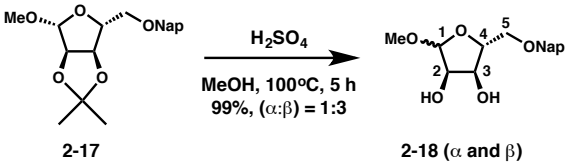
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.88 – 7.80 (m, 3H, Nap), 7.78 (s, 1H, Nap), 7.51 – 7.44 (m, 3H, Nap), 4.97 (s, 1H, H-1), 4.72 (s, 2H, -OC*H*$_2$Nap), 4.70 (d, *J* = 6.1 Hz, 1H, H-2), 4.57 (d, *J* = 6.0 Hz, 1H, H-3), 4.41 (ddd, *J* = 7.7, 6.5, 1.1 Hz, 1H, H-4), 3.56 (dd, *J* = 9.7, 6.4 Hz, 1H, H-5a), 3.50 (dd, *J* = 9.7, 8.1 Hz, 1H, H-5b), 3.29 (s, 3H, -OCH_3), 1.49 (s, 3H, -CH_3), 1.32 (s, 3H, -CH_3).

^{13}C NMR (126 MHz, $CDCl_3$) δ 135.6, 133.3, 133.1, 128.3, 128.0, 127.8, 126.5, 126.2, 126.0, 125.8, 112.5, 109.4, 85.3, 85.3, 82.3, 73.4, 71.2, 54.9, 26.6, 25.1.

HRMS (ESI) m/*z* calcd for $C_{20}H_{25}O_5$ ($[M+H]^+$) 345.1702, found 345.1695.

IR (neat) ν 3052, 2996, 2938, 1601, 1510, 1458, 1372, 1271, 1243, 1210, 1107, 1059, 1018, 966, 870, 818, 753 cm^{-1}.

[a]$_D^{22}$ -41.0 (*c* = 10.0, $CHCl_3$)



2-17 **2-18 (α and β)**

Compound 2-18: To a 1 L round bottom flask was added compound **2-17** (12.16 g, 35.3 mmol, 1 eq) and MeOH (500 mL). The flask was placed in an oil bath at 114 °C and equipped with a distillation head. While the solution was warming to the bath temperature, concentrated H_2SO_4 (4.6 mL) was added slowly over 5 minutes. The solution was heated to a vigorous boil and the distillate was collected. Fresh anhydrous methanol was slowly added to the solution to replenish the methanol that was lost. After 5 h the solution was cooled to room temperature. The reaction was quenched with solid $NaHCO_3$ (20 g), the solution was decanted and concentrated. The organic slurry was diluted with CH_2Cl_2, filtered, and the filtrate was once again concentrated. The crude oil was purified by silica gel chromatography to furnish compound **2-18** as a white solid (10.68 g, 99%, α:β = ~1:3).

α-anomer (R_f 0.23, 75:25 ethyl acetate:hexanes)

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.86 – 7.80 (m, 3H, Nap), 7.75 (s, 1H, Nap), 7.53 – 7.40 (m, 3H, Nap), 4.96 (d, *J* = 4.5 Hz, 1H, H-1), 4.74 (d, *J* = 12.1 Hz, 1H, C*H*$_2$Nap), 4.69 (d, *J* = 12.2 Hz, 1H, C*H*$_2$Nap), 4.21 – 4.12 (m, 2H, H-2 + H-3), 3.98 (ddd, *J* = 7.7, 6.1, 3.1 Hz, 1H, H-4), 3.65 (dd, *J* = 10.5, 3.4 Hz, 1H, H-5a), 3.62 (dd, *J* = 10.5, 4.3 Hz, 1H, H-5b), 3.49 (s, 3H, -OCH_3), 2.92 (d, *J* = 9.1 Hz, 1H, -OH), 2.66 (d, *J* = 8.1 Hz, 1H, -OH).

^{13}C NMR (126 MHz, $CDCl_3$) δ 135.4, 133.3, 133.1, 128.4, 128.0, 127.8, 126.7, 126.3, 126.1, 125.8, 103.0, 84.0, 73.8, 71.7, 71.5, 70.2, 55.8.

HRMS (ESI) m/*z* calcd for $C_{17}H_{20}O_5Na$ ($[M+Na]^+$) 327.1208, found 327.1210.

IR (neat) ν 3410, 3059, 2926, 2850, 1635, 1507, 1465, 1448, 1403, 1372, 1347, 1271, 1188, 1125, 1094, 1039, 955, 896, 855, 817, 758 cm^{-1}.

$[a]_D^{22}$ + 72.5 (*c* = 1.9, $CHCl_3$)

β-anomer (R_f 0.34, 75:25 ethyl acetae:hexanes)

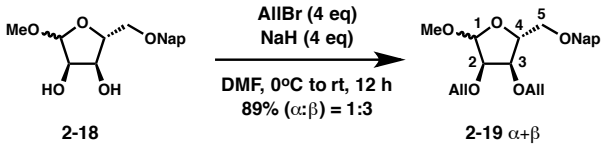
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.81 (td, *J* = 6.0, 5.5, 3.3 Hz, 3H, Nap), 7.77 (s, 1H, Nap), 7.51 – 7.44 (m, 3H, Nap), 4.83 (s, 1H, H-1), 4.72 (d, *J* = 12.3 Hz, 1H, C*H*$_2$Nap), 4.69 (d, *J* = 12.3 Hz, 1H, C*H*$_2$Nap), 4.23 – 4.09 (m, 3H, H-2 + H-3 + H-4), 3.99 (d, *J* = 4.6 Hz, 1H, -OH), 3.93 (s, 1H, -OH), 3.65 (dd, *J* = 10.3, 4.0 Hz, 1H, H-5a), 3.55 (dd, *J* = 10.3, 6.5 Hz, 1H, H-5b), 3.29 (s, 3H, -OCH_3).

^{13}C NMR (126 MHz, $CDCl_3$) δ 135.3, 133.2, 133.0, 128.2, 127.9, 127.7, 126.6, 126.1, 125.9, 125.8, 108.2, 81.9, 74.8, 73.5, 72.5, 72.1, 55.1.

HRMS (ESI) m/*z* calcd for $C_{17}H_{20}O_5Na$ ($[M+Na]^+$) 327.1208, found 327.1199.

IR (neat) ν 3390, 3045, 1639, 1604, 1503, 1451, 1354, 1264, 1188, 1113, 1063, 1028, 976, 945, 886, 858, 817, 752 cm^{-1}.

$[a]_D^{22}$ -32.5 (*c* = 4.55, $CHCl_3$)



2-18 **2-19** α+β

Compound 2-19: To a flame dried 500 mL round bottom flask was added NaH (3.37 g, 140.0 mmol, 4 eq) followed by DMF (80 mL) and the mixture was cooled to 0 ^oC. Then, compound **2-18** (10.68 g, 35.1 mmol, 1 eq) was added dropwise as a solution in DMF (80 mL). After 30 min at 0 ^oC, allyl bromide (11.9 mL, 140.0 mmol, 4 eq) was added, the reaction was slowly warmed to room temperature and was stirred for 12 h. The reaction was cooled to 0 ^oC, quenched with a saturated aqueous NH_4Cl solution, diluted with H_2O and extracted with EtOAc three times. The

organic layer was dried over Na_2SO_4, concentrated, and purified by silica gel chromatography. Compound **2-19** was obtained as a slightly yellow oil (11.985 g, 89%, α:β = ~1:3).

α-anomer (R_f 0.21, 30:70 ethyl acetate:hexanes)

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.87 – 7.78 (m, 3H, Nap), 7.76 (s, 1H, Nap), 7.51 – 7.43 (m, 3H, Nap), 5.97 (ddt, *J* = 17.2, 10.3, 5.9 Hz, 1H, -CH_2CH=CH_2), 5.87 (ddt, *J* = 17.3, 10.3, 5.9 Hz, 1H, -CH_2CH=CH_2), 5.28 (dd, *J* = 17.2, 1.6 Hz, 1H, -CH_2CH=CH_2), 5.22 – 5.13 (m, 2H, -CH_2CH=CH_2), 5.09 (dd, *J* = 10.3, 1.7 Hz, 1H, -CH_2CH=CH_2), 4.97 (d, *J* = 4.3 Hz, 1H, H-1), 4.77 (d, *J* = 12.3 Hz, 1H, -CH_2Nap), 4.69 (d, *J* = 12.3 Hz, 1H, -CH_2Nap), 4.26 (appq, *J* = 4.0 Hz, 1H, H-4), 4.17 – 4.10 (m, 3H, -CH_2CH=CH_2), 4.06 (dd, *J* = 13.2, 5.9 Hz, 1H, -CH_2CH=CH_2), 3.89 (dd, *J* = 6.7, 2.9 Hz, 1H, H-3), 3.84 (dd, *J* = 6.8, 4.3 Hz, 1H, H-2), 3.60 (d, *J* = 4.0 Hz, 2H, H-5a + H-5b), 3.46 (s, 3H, -OCH_3).

^{13}C NMR (126 MHz, $CDCl_3$) δ 135.5, 135.0, 134.7, 133.3, 133.1, 128.3, 127.9, 127.8, 126.6, 126.3, 126.0, 125.7, 117.9, 117.6, 102.5, 82.2, 78.2, 75.5, 73.7, 72.0, 71.9, 70.3, 55.5.

HRMS (ESI) m/*z* calcd for $C_{23}H_{28}O_5Na$ ($[M+Na]^+$) 407.1834, found 407.1830

IR (neat) ν 2912, 1642, 1594, 1509, 1346, 1275, 1108, 1040, 910, 856, 818, 753 cm^{-1}**.**

[a]$_D^{22}$ + 60.1 (*c* = 5.8, $CHCl_3$)

β-anomer (R_f 0.58, 30:70 ethyl acetate:hexanes)

^{1}H NMR (500 MHz, $CDCl_3$) d 7.87 – 7.75 (m, 4H, Nap), 7.52 – 7.41 (m, 3H, Nap), 5.99 – 5.83 (m, 2H, -CH_2CH=CH_2), 5.31 (dd, *J* = 17.2, 1.6 Hz, 1H, -CH_2CH=CH_2), 5.25 (dd, *J* = 17.2, 1.6 Hz, 1H, -CH_2CH=CH_2), 5.20 (dd, *J* = 10.4, 1.5 Hz, 1H, -CH_2CH=CH_2), 5.16 (dd, *J* = 10.4, 1.5 Hz, 1H, -CH_2CH=CH_2), 4.92 (s, 1H, H-1), 4.78 (d, *J* = 12.6 Hz, 1H, -CH_2Nap), 4.74 (d, *J* = 12.2 Hz, 1H, -CH_2Nap), 4.32 (ddd, 7.07, 5.85, 3.59 Hz, 1H, H-4), 4.19 – 4.04 (m, 3H, -CH_2CH=CH_2),

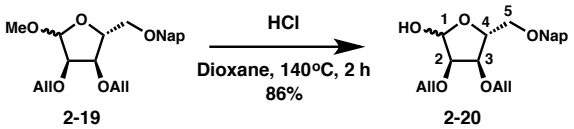
4.03 – 3.97 (m, 2H, -C*H*$_2$CH=CH$_2$ + H-3), 3.83 (d, *J* = 4.6 Hz, 1H, H-2), 3.69 (dd, *J* = 10.6, 3.6 Hz, 1H, H-5a), 3.59 (dd, *J* = 10.6, 5.9 Hz, 1H, H-5b), 3.35 (s, 3H, -OCH$_3$).

^{13}C NMR (126 MHz, CDCl$_3$) d 135.9, 134.5, 134.4, 133.3, 133.0, 128.1, 127.9, 127.8, 126.3, 126.1, 125.9, 125.7, 117.7, 117.6, 106.5, 80.5, 79.7, 78.4, 73.3, 71.6, 71.5, 71.5, 55.2.

HRMS (ESI) m/*z* calcd for $C_{23}H_{28}O_5Na$ ([M+Na]$^+$) 407.1834, found 407.1839.

IR (neat) ν 3059, 2916, 2850, 1639, 1507, 1458, 1354, 1254, 1106, 1066, 1039, 983, 935, 818, 754 cm^{-1}.

[a]$_D^{21}$ + 7.3 (*c* = 4.97, CHCl$_3$)



2-19 2-20

Compound 2-20: To a 1 L round bottom flask was added compound **2-19** (11.95 g, 31.1 mmol, 1 eq) and 1,4-dioxane (150 mL). The flask was fitted with an air condenser and a rubber septum without positive pressure of nitrogen and heated in an oil bath. While the solution was warming, 1 M HCl (150 mL) was added in portions over 5 min. The solution was heated at 140 ^{o}C for 2 h. The reaction was quenched with saturated aqueous NaHCO$_3$ and extracted with EtOAc three times. The organic layer was dried over NaSO$_4$, concentrated, and purified by silica gel chromatography (R$_f$=0.6 in 50:50 hexanes:ethyl acetate) to provide compound **2-20** as a yellow oil (9.91 g, 86%, 1.6:1 anomeric mixture).

^{1}H NMR (500 MHz, CDCl$_3$) δ 7.84 –7.71 (m, 4H major, 4H minor, Nap), 7.51 – 7.39 (m, 3H major, 3H minor, Nap), 5.99 – 5.80 (m, 2H major, 2H minor, -CH$_2$C*H*=CH$_2$), 5.34 – 5.12 (m, 5H major, 5H minor, H-1 + -CH$_2$CH=C*H*$_2$), 4.83 – 4.62 (m, 2H major, 2H minor, -C*H*$_2$Nap), 4.33 (ddd, *J* = 4.0, 3.9, 2.4 Hz, 1H major, H-4), 4.26 (ddd, *J* = 6.6, 3.2, 3.2 Hz, 1H minor, H-4), 4.21 – 3.90 (m, 6H major, 5H minor, H-3 (major + minor), H-2 (major), -C*H*$_2$CH=CH$_2$), 3.83 (d, *J* = 4.7

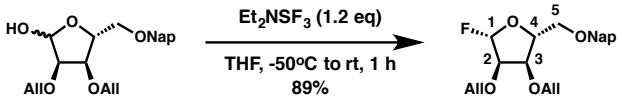
Hz, 1H minor, H-2), 3.72 (dd, *J* = 10.4, 2.9 Hz, 1H minor, H-5a), 3.61 (dd, *J* = 10.4, 3.5 Hz, 1H minor, H-5b), 3.57 (dd, *J* = 10.5, 3.7 Hz, 1H major, H-5a), 3.53 (dd, *J* = 10.5, 4.4 Hz, 1H major, H-5b).

^{13}C NMR (100 MHz, $CDCl_3$) δ 135.4 (major), 135.0 (minor), 134.4 (minor), 134.4 (minor), 134.2 (major), 134.1 (major), 133.3 (minor), 133.3 (major), 133.1 (minor), 133.1 (major), 128.4 (minor), 128.3 (major), 127.9 (minor), 127.9 (major), 127.8 (major), 127.8 (minor), 126.8 (minor), 126.5 (major), 126.3 (major), 126.3 (minor), 126.1 (minor), 126.0 (major), 125.8 (minor), 125.6 (major), 117.8 (major), 117.7 (minor), 117.7 (minor), 117.6 (major), 100.5 (minor, C-1), 96.2 (major, C-1), 80.9 (minor, C-4), 80.7 (major, C-4), 77.9 (minor, C-2), 77.7 (major), 77.5 (minor), 77.4 (minor, C-3), 73.7 (major), 73.6 (minor), 71.9 (major), 71.7 (major), 71.6 (minor), 71.5 (minor), 70.1 (major), 70.0 (minor).

HRMS (ESI) m/*z* calcd for $C_{22}H_{26}O_5Na$ ($[M+Na]^+$) 393.1678, found 393.1674

IR (neat) ν 3421, 3045, 2912, 2863, 1646, 1509, 1455, 1427, 1344, 1271, 1089, 1032, 924, 856, 817, 753 cm^{-1}.

$[a]_D^{21}$ +37.5 (*c* = 8.5, $CHCl_3$)



2-20 2-6

Compound 2-6: To a 25 mL round-bottomed flask was added compound **2-20** (5.0 g, 13.5 mmol, 1 eq) and THF (65 mL). The solution was cooled to -50 °C and diethylaminosulfur trifluoride (2.1 mL, 16.2 mmol, 1.2 eq) was added slowly over 2 min. The solution was immediately warmed to room temperature and stirred for 1 h. The solution was cooled to -50 °C and quenched with saturated aqueous $NaHCO_3$ (20 mL). The solution was warmed to room temperature, the organic layer was extracted with EtOAc three times, dried over Na_2SO_4,

concentrated, and purified by silica gel chromatography to give compound **2-6** as a pale yellow oil (4.48 g, 89%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.87 – 7.77 (m, 4H, Nap), 7.51 – 7.44 (m, 3H, Nap), 5.99 – 5.81 (m, 2H, -CH_2CH=CH_2), 5.73 (d, *J* = 63.4 Hz, 1H, H-1), 5.37 – 5.15 (m, 4H, -CH_2CH=CH_2), 4.81 (d, *J* = 12.3 Hz, 1H, CH_2Nap), 4.75 (d, *J* = 12.3 Hz, 1H, CH_2Nap), 4.39 (dddd, *J* = 8.2, 8.2, 5.3, 3.2 Hz, 1H, H-4), 4.18 (appdt, *J* = 5.7, 1.4 Hz, 1H, H-3), 4.14 – 3.97 (m, 5H, -CH_2CH=CH_2, + H-2), 3.77 (dd, *J* = 11.1, 3.2 Hz, 1H, H-5a), 3.67 (dd, *J* = 11.0, 5.4 Hz, 1H, H-5b).

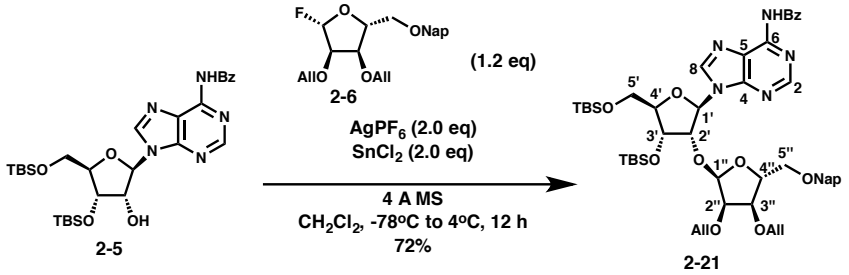
^{13}C NMR (125 MHz, $CDCl_3$) δ 135.7, 134.2, 134.1, 133.4, 133.1, 128.3, 128.0, 127.8, 126.5, 126.2, 126.0, 125.8, 118.3, 118.0, 112.8 (d, *J* = 224.4 Hz), 82.5 (d, *J* = 2.6 Hz), 79.1 (d, *J* = 29.9 Hz), 77.1, 73.6, 72.1, 72.0, 70.4.

^{19}F NMR (470 MHz, $CDCl_3$) δ -118.4 (d, *J* = 63.6 Hz).

HRMS (ESI) m/*z* calcd for $C_{22}H_{25}O_4NaF$ ($[M+Na]^+$) 395.1635, found 395.1638

IR (neat) ν 3052, 2919, 2862, 1646, 1601, 1510, 1469, 1420, 1347, 1271, 1101, 994, 948, 855, 818, 747 cm^{-1}.

$[a]_D^{21}$ +46.6 (c = 4.8, $CHCl_3$)



Compound 2-21: To a 250 mL schlenk flask was added $SnCl_2$ (3.39 g, 17.9 mmol, 2.0 eq) and freshly dried powdered 4 Å molecular sieves (6.0 g). Dried $AgPF_6$ (4.53 g, 17.90 mmol, 2.0 eq), contained in a 7 mL vial, was added to the flask keeping the $AgPF_6$ and $SnCl_2$ segregated. The flask was dried overnight under high vacuum. To the flask was added CH_2Cl_2 (60 mL) and the

reaction was cooled to -78 ^{o}C. Lastly, compound **2-6** (4.0 g, 10.7 mmol, 1.2 eq) and compound **2-5** (5.37 g, 8.95 mmol, 1 eq) were added in CH_2Cl_2 (40 mL). Compounds **2-6** and **2-5** had been previously dried by azeotropic removal of water with dry CH_3CN three times on a rotary evaporator filled with dry nitrogen and dried over P_2O_5 under high vacuum for 5 h. The solution was stirred for 30 min at -78 ^{o}C and then stirred at 4 ^{o}C for 15 h. The solution was quenched with a 0^{o}C saturated aqueous $NaHCO_3$ solution and stirred for 3 h, after which point the biphasic mixture was filtered through celite washing with CH_2Cl_2. The organic layer was dried over Na_2SO_4, concentrated, and purified by silica gel chromatography (R_f 0.5 in 50:50 hexanes:ethyl acetate) to give compound **2-21** as a pale yellow foam (6.04 g, 72%).

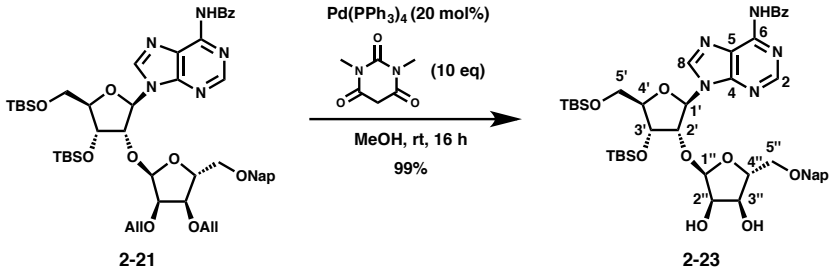
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.93 (s, 1H, N*H*Bz), 8.81 (s, 1H, H-8), 8.39 (s, 1H, H-2), 8.00 (d, *J* = 7.0 Hz, 2H, Bz), 7.86 – 7.75 (m, 3H, Nap), 7.71 (s, 1H, Nap), 7.61 (t, *J* = 7.5 Hz, 1H, Bz), 7.52 (t, *J* = 7.5 Hz, 2H, Bz), 7.48 – 7.42 (m, 2H, Nap), 7.39 (dd, *J* = 8.5, 1.7 Hz, 1H, Nap), 6.33 (d, *J* = 5.0 Hz, 1H, H-1'), 5.89-5.78 (m, 2H, -CH_2CH=CH_2), 5.23 – 5.10 (m, 5H, -CH_2CH=CH_2 + H-1''), 4.84 (appt, *J* = 4.8 Hz, 1H, H-2'), 4.69 (d, *J* = 12.3 Hz, 1H, -CH_2Nap), 4.61 (d, *J* = 12.3 Hz, 1H, -CH_2Nap), 4.53 (appt, *J* = 4.4 Hz, 1H, H-3'), 4.23-4.17 (m, 2H, H-4' + H-4''), 4.13 – 4.02 (m, 2H, - CH_2CH=CH_2), 4.02 – 3.91 (m, 3H, -CH_2CH=CH_2 + H-5'a), 3.85 (dd, *J* = 6.4, 4.4 Hz, 1H, H-3''), 3.81 – 3.72 (m, 2H, H-2''+ H-5'b), 3.52 (d, *J* = 3.7 Hz, 2H, H-5''a + H-5''b) 0.93 (s, 9H, -tBu), 0.91 (s, 9H, -tBu), 0.16 (s, 3H, -CH_3), 0.14 (s, 3H, -CH_3), 0.10 (s, 3H, -CH_3), 0.10 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 164.7, 152.8, 151.6, 149.5, 142.3, 135.6, 134.9, 134.6, 133.9, 133.3, 133.1, 132.8, 129.0, 128.3, 127.9, 127.9, 127.8, 126.5, 126.3, 126.0, 125.7, 123.3, 117.6, 117.5, 101.5, 87.0, 85.6, 81.9, 79.6, 77.8, 75.6, 73.7, 72.0, 71.6, 71.2, 69.9, 62.3, 26.2, 26.0, 18.6, 18.3, -4.4, -5.0, -5.3, -5.3.

HRMS (ESI) m/*z* calcd for $C_{51}H_{70}N_5O_9Si_2$ ([M+H]$^+$) 952.4712, found 952.4711.

IR (neat) ν 2933, 2857, 1698, 1611, 1580, 1455, 1250, 1084, 1039, 834, 782 cm^{-1}.

[a]$_D^{21}$ +2.7 (c = 2.41, $CHCl_3$)



Compound 2-23: To a 125 mL round bottomed flask containing compound **2-21** (1.0 g, 1.05 mmol, 1 eq) was added $Pd(PPh_3)_4$ (240 mg, 0.21 mmol, 0.2 eq) and 1,3-Dimethylbarbituric acid (1.62 g, 10.3 mmol, 10 eq). Methanol (30 mL) was added and the reaction was stirred at room temperature for 16 h. The reaction was quenched at 0^{o}C with an aqueous Na_2CO_3 solution and stirred for 2 h at room temperature. The reaction mixture was extracted three times with ethyl acetate and three times with chloroform, filtered through Na_2SO_4, and concentrated. The product was purified by silica gel chromatography (R_f 0.53 in 40:60 hexanes:ethyl acetate) to yield compound **2-23** as a bright yellow foam (917 mg, 99%).

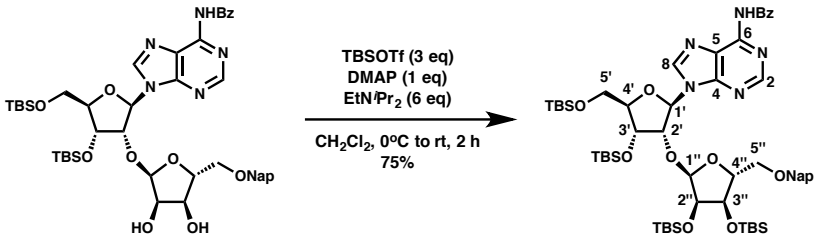
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.92 (s, 1H, N*H*Bz), 8.62 (s, 1H, H-8), 8.40 (s, 1H, H-2), 8.02 (dd, J=7.0 Hz, 1.5 Hz, 2 H, -Bz), 7.83-7.80 (m, 3H, -Nap), 7.69-7.44 (m, 7H, -Nap+Bz), 6.25 (d, *J* = 2.7 Hz, 1H, H-1'), 5.37 (d, *J* = 4.2 Hz, 1H, H-1''), 4.73-4.63 (m, 4H, CH_2Nap + H-2' + H-3;), 4.26-4.25 (m, 1H, H-4''), 4.21 (ddd, *J* = 9.8, 5.7, 4.1 Hz, 1H, H-2''), 4.13 (dd, *J* = 6.0, 2.8 Hz, 1H, H-4'), 4.09-3.98 (m, 2H, H-5'a, H-3''), 3.81 (dd, *J* = 11.7, 2.4 Hz, 1H, H-5'b), 3.62 (dd, *J* = 10.6, 3.8 Hz, 1H, H-5''a), 3.58 (dd, J = 10.6, 3.6 Hz, 1H, H-5''b), 3.06 (d, *J* = 9.8 Hz, 1H, -OH), 2.94 (d, *J* = 10.7 Hz, 1H, -OH), 0.94 (s, 9H, -tBu), 0.92 (s, 9H, -tBu), 0.151 (s, 6H, -CH_3), 0.11 (s, 3H, -CH_3), 0.10 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 164.7, 152.9, 151.2, 149.6, 141.4, 135.3, 133.8, 133.3, 133.1, 132.9, 129.0, 128.4, 128.0, 127.9, 127.8, 126.6, 126.3, 126.1, 125.7, 123.4, 101.5, 87.3, 85.1, 84.6, 79.5, 73.8, 72.0, 72.0, 67.0, 69.8, 61.3, 26.2, 26.0, 18.6, 18.3, -4.1, -4.7, -5.2, -5.3.

HRMS m/z calcd for $C_{45}H_{62}N_5O_9Si_2$ ($[M+H]^+$) 872.4086, found 872.4077

IR (neat) ν 3344, 2926, 2850, 1698, 1608, 1576, 1458, 1438, 1254, 1119, 830 cm^{-1}.

$[a_D]^{21}$ +15.3 (c = 15.3, $CHCl_3$)



2-23 2-24

Compound 2-24: A 7 mL reaction vial containing compound **2-23** (910 mg, 1.043 mmol, 1 eq) and 4-dimethylaminopyridine (127 mg, 1.043 mmol, 1 eq) was evacuated and backfilled three times with dry nitrogen. Methylene chloride (10 mL) was added and the vial was cooled to 0 ºC. Next, EtN^iPr_2 (1.09 mL, 6.26 mmol, 6 eq) and TBSOTf (0.72 mL, 3.13 mmol, 3 eq) were added and the reaction was allowed to warm to room temperature. A saturated aqueous NH_4Cl solution was added after 2 h, and the reaction was stirred vigorously for 1 h. The product was purified by silica gel chromatography (R_f 0.68 in 60:40 hexanes:ethyl acetate) to yield compound **2-24** (862 mg, 75%) as a white foam.

^{1}H NMR (500 MHz, $CDCl_3$) d 8.93 (s, 1H, N*H*Bz), 8.78 (s, 1H, H-8), 8.39 (s, 1H, H-2), 8.01 (d, J=7.1 Hz, 2H, -Bz), 7.85 – 7.76 (m, 3H, -Nap), 7.74 (s, 1H, -Nap), 7.61 (t, J=7.6 Hz, 1H, -Bz), 7.53 (t, *J* = 7.6 Hz, 2H, -Bz), 7.49 – 7.45 (m, 2H, Nap), 7.42 (dd, *J* = 8.4, 1.7 Hz, 1H, -Nap), 6.23 (d, *J* = 3.2 Hz, 1H, H-1'), 5.28 (d, *J* = 3.7 Hz, 1H, H-1''), 4.76 – 4.62 (m, 3H, -C*H*$_2$Nap + H-2'), 4.55 (dd, *J* = 6.1, 4.5 Hz, 1H, H-3'), 4.24-4.18 (m, 2H, H-4' + H-4''), 4.05-3.99 (m, 3H, H-2'' +

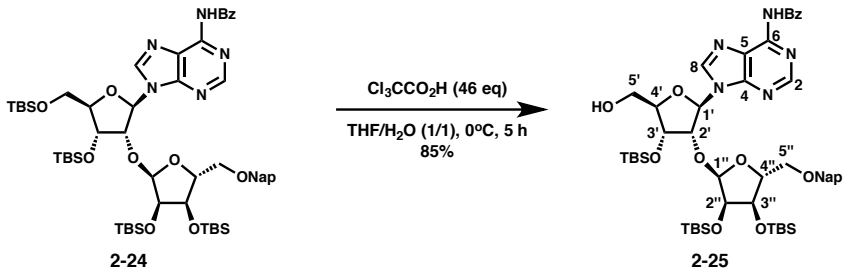
H-3'' + H-5'a), 3.78 (dd, *J* = 11.5, 2.5 Hz, 1H, H-5'b), 3.62 (dd, *J* = 10.7, 3.0 Hz, 1H, H-5''a), 3.57 (dd, *J* = 10.8, 3.8 Hz, 1H, H-5''b), 0.91 (s, 18H, -tBu), 0.86 (s, 9H, -tBu), 0.85 (s, 9H, -tBu), 0.13 (s, 3H, $-CH_3$), 0.10 (s, 6H, $-CH_3$), 0.09 (s, 3H, $-CH_3$), 0.03 (s, 6H, $-CH_3$), 0.00 (s, 3H, $-CH_3$), -0.01 (s, 3H, $-CH_3$).

^{13}C NMR (125 MHz, $CDCl_3$) d 164.7, 152.8, 151.3, 149.4, 142.1, 135.8, 134.0, 133.4, 133.1, 132.8, 129.0, 128.2, 127.9, 127.9, 127.8, 126.4, 126.2, 126.0, 125.7, 123.4, 103.2, 87.9, 84.6, 83.6, 81.0, 73.7, 73.6, 71.5, 70.0, 69.7, 61.7, 26.2, 26.2, 26.1, 26.0, 18.6, 18.5, 18.3, 18.3, -4.2, -4.3, -4.4, -4.5, -4.6, -4.8, -5.2, -5.2.

HRMS m/z calcd for $C_{57}H_{90}N_5O_9Si_4$ ($[M+H]^+$) 1100.5816, found 1100.5829

IR (neat) n 2928, 2857, 1701, 1610, 1576, 1455, 1252, 1070, 836, 778 cm^{-1}

$[a_D]^{23}$ + 10.1 (c = 1.04, $CHCl_3$)



Compound 2-25: To a 100 mL round bottomed flask was added compound **2-24 (**530 mg, 0.48 mmol, 1 eq) and THF (12 mL). The solution was cooled to 0 ^oC. A cold solution of trichloroacetic acid (3.5 g, 21.7 mmol, 45 eq) in water (12 mL) was added dropwise at 0 ^oC over 30 min. The reaction was stirred at 0 ^oC for 5 h, at which point it was quenched by pouring into a 0 ^oC solution of saturated aqueous $NaHCO_3$. The quenched reaction was extracted three times with EtOAc, dried through Na_2SO_4 and concentrated. Silica gel chromatography (R_f=0.46 in 60:40 hexanes:ethyl acetate) yielded compound **2-25** as a white foam (403 mg, 85%).

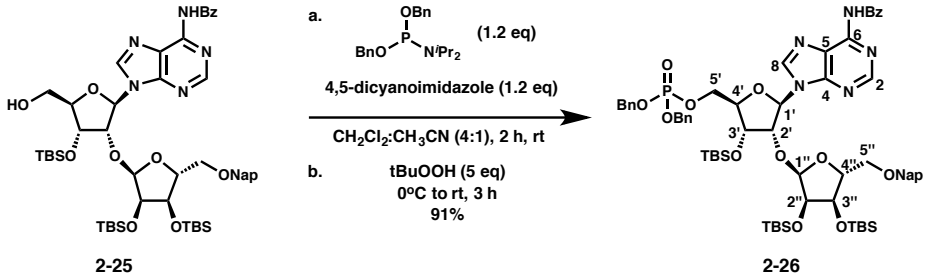
^{1}H NMR (500 MHz, $CDCl_3$) δ 9.19 (s, 1H, -N*H*Bz), 8.78 (s, 1H, H-8), 8.07 (s, 1H, H-2), 8.00 (d, J=7.6 Hz, 2H, -Bz), 7.84 – 7.78 (m, 3H, -Nap), 7.75 (s, 1H, -Nap), 7.59 (t, *J* = 7.6 Hz, 1H, -Bz), 7.51-7.41 (m, 5H, -Nap + -Bz), 6.07 (d, *J* = 7.5 Hz, 1H, H-1'), 5.84 (bd, J=10.6 Hz, 1H, -OH), 4.99 (dd, *J* = 7.5, 4.4 Hz, 1H, H-2'), 4.89 (d, *J* = 3.6 Hz, 1H, H-1''), 4.67 (d, *J* = 3.5 Hz, 2H, -CH_2Nap), 4.62 (dd, *J* = 4.5, 1.3 Hz, 1H, H-3'), 4.24 – 4.16 (m, 2H, H-4' + H-4''), 4.02 – 3.92 (m, 2H, H-5'a + H-3''), 3.85 (dd, *J* = 4.9, 3.6 Hz, 1H, H-2''), 3.76 (ddd, *J* = 12.9, 11.1, 1.9 Hz, 1H, H-5'b), 3.59 (dd, *J* = 10.8, 3.0 Hz, 1H, H-5''a), 3.53 (dd, *J* = 10.8, 3.6 Hz, 1H, H-5''b), 0.93 (s, 9H, -tBu), 0.86 (s, 9H, -tBu), 0.75 (s, 9H, -tBu), 0.14 (s, 3H, -CH_3), 0.10 (s, 3H, -CH_3), -0.01 (s, 3H, -CH_3), -0.02 (s, 3H, -CH_3), -0.09 (s, 3H, -CH_3), -0.12 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 164.5, 152.2, 150.6, 150.3, 143.7, 135.7, 133.6, 133.3, 133.0, 132.9, 129.0, 128.2, 127.9, 127.9, 127.7, 126.4, 126.1, 125.9, 125.8, 124.5, 104.1, 89.3, 89.2, 83.0, 81.4, 73.6, 73.6, 73.0, 71.4, 69.4, 62.7, 26.1, 26.1, 26.1, 18.5, 18.2, 18.2, -4.2, -4.3, -4.4, -4.5, -4.6, -4.7.

HRMS (ESI) m/z calcd for $C_{51}H_{76}N_5O_9Si_3$ ($[M+H]^+$) 986.4951, found 986.4959

IR (neat) ν 3254, 2919, 2857, 1705, 1611, 1580, 1458, 1251, 1153, 1108, 834, 775 cm^{-1}

$[a_D]^{23}$ +3.0 (*c* = 1.08, $CHCl_3$)



2-25 2-26

Compound 2-26: To a dry 20 mL reaction vial was added compound **2-25** (520 mg, 0.47 mmol, 1 eq) and dibenzyl N,N-diisopropylphosphoramidite (195 mg, 0.56 mmol, 1.2 eq). The compounds were co-evaporated three times on a N_2-filled rotary evaporator with dry CH_3CN and

then dried over P_2O_5 overnight. Separately, 4,5-dicyanoimidazole (66.6 mg, 0.56 mmol, 1.2 eq) was evaporated and dried in a similar manner. To compound **2-21** and the phosphoramidite was added dry CH_2Cl_2 (4.7 mL), followed by a solution of the 4,5-dicyanoimidazole in CH_3CN (1.5 mL) at room temperature. After 2 h, the reaction was cooled to 0 ^oC and a solution of t-butyl peroxide (0.39 mL of a 5.5 M solution in decane, 2.35 mmol, 5 eq) was added. The reaction was warmed to room temperature and stirred for 3 h. Next, the reaction was poured into a saturated aqueous $NaHCO_3$ solution, extracted three times with ethyl acetate, dried through Na_2SO_4, and concentrated. The product was purified by silica gel chromatography (R_f 0.29 in 50:50 hexanes:ethyl acetate) to give compound **2-26** as a colorless oil (533 mg, 91%).

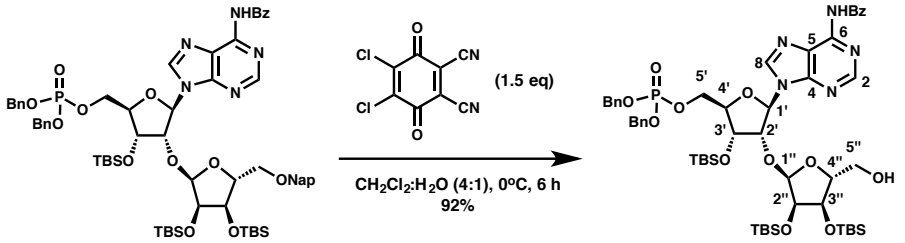
^{1}H NMR (500 MHz, $CDCl_3$) δ 9.07 (s, 1H, -N*H*Bz), 8.74 (s, 1H, H-8), 8.27 (s, 1H, H-2), 8.01 (d, *J* = 7.0 Hz, 2H, -Bz), 7.86 – 7.77 (m, 3H, -Nap), 7.75 (s, 1H, -Nap), 7.60 (t, *J* = 7.4 Hz, 1H, -Bz), 7.55 – 7.39 (m, 5H, -Nap + -Bz), 7.33 – 7.27 (m, 10H), 6.18 (d, *J* = 3.1 Hz, 1H, H-1'), 5.26 (d, *J* = 3.8 Hz, 1H, H-1''), 5.00 (appt, *J* = 8.1 Hz, 4H, -C*H*$_2$Bn), 4.83 (dd, J = 4.4, 3.5 Hz, 1H, H-2'), 4.72 (d, *J* = 12.0 Hz, 1H, -C*H*$_2$Nap), 4.66 (d, *J* = 12.2 Hz, 1H, -C*H*$_2$Nap), 4.55 (dd, *J* = 6.2, 4.5 Hz, 1H, H-3'), 4.40 (ddd, *J* = 11.2, 5.6, 3.6 Hz, 1H, H-5'a), 4.35 – 4.28 (m, 1H, H-4'), 4.25 (appq, *J* = 3.6 Hz, 1H, H-4''), 4.19 (ddd, *J* = 11.2, 5.2, 3.7 Hz, 1H, H-5'b), 4.03 (dd, J = 5.3, 4.3 Hz, 1H, H-3''), 4.00 (dd, *J* = 5.3, 3.9 Hz, 1H, H-2''), 3.63 (dd, *J* = 10.8, 3.1 Hz, 1H, H-5''a), 3.58 (dd, *J* = 10.8, 3.7 Hz, 1H, H-5''b), 0.91 (s, 9H, -tBu), 0.88 (s, 9H, -tBu), 0.83 (s, 9H, -tBu), 0.13 (s, 3H, -CH_3), 0.09 (s, 3H, -CH_3), 0.04 (s, 3H, -CH_3), -0.00 (s, 6H, -CH_3), -0.02 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 164.5, 152.7, 151.2, 149.6, 142.2, 135.7, 135.6, 133.8, 133.3, 133.1, 132.8, 128.9, 128.6, 128.6, 128.2, 128.0, 128.0, 128.0, 127.8, 127.8, 126.4, 126.2, 125.9,

125.7, 123.7, 103.5, 88.5, 83.9, 82.3 (d, *J* = 8.7 Hz), 80.3, 73.6, 71.5, 70.4, 69.7, 69.5 (appt, *J* = 4.9 Hz), 65.5 (d, *J* = 3.9 Hz), 26.1, 26.1, 25.9, 18.4, 18.2, 18.2, -4.3, -4.4, -4.5, -4.6, -4.6, -4.9.

^{31}P NMR (202 MHz, $CDCl_3$) δ 0.030

HRMS (ESI) m/z calcd for $C_{65}H_{89}N_5O_{12}Si_3P$ ($[M+H]^+$) 1246.5553, found 1246.5544



2-26 2-27

Compound 2-27: To a 20 mL reaction vial was added compound **2-26** (308 mg, 0.247 mmol, 1 eq), methylene chloride (5.3 mL) and deionized water (1.2 mL). The solution was cooled to 0 °C and 2,3-Dichloro-5,6-dicyano-1,4-benzoquinone (84.1 mg, 0.371 mmol, 1.5 eq) was added. The reaction was stirred at 0 °C for 2 h. After 2 h, the reaction was poured into a saturated aqueous $Na_2S_2O_3$ solution, extracted three times with ethyl acetate, dried through Na_2SO_4, and concentrated. The product was purified by silica column chromatography (R_f 0.69 in 30:70 hexanes:ethyl acetate) to yield compound **2-27** as a colorless oil (251 mg, 92%).

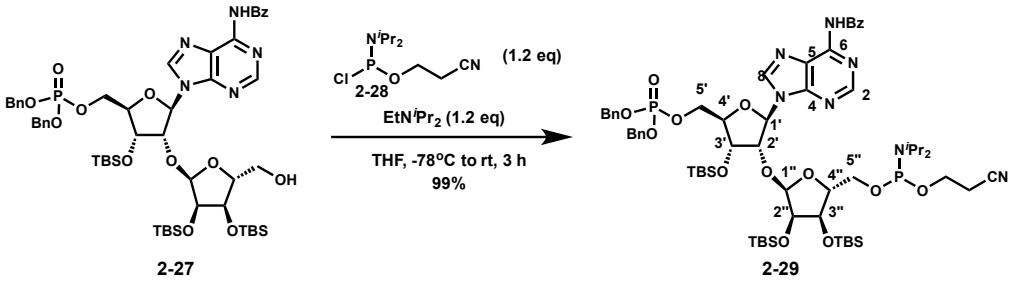
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.97 (s, 1H, -N*H*Bz), 8.77 (s, 1H, H-8), 8.25 (s, 1H, H-2), 8.01 (d, *J* = 7.4, 2H, -Bz), 7.62 (t, *J* = 7.4 Hz, 1H, -Bz), 7.54 (t, *J* = 7.7 Hz, 2H, -Bz), 7.31-7.29 (m, 10H, -Bn), 6.16 (d, *J* = 4.1 Hz, 1H, H-1'), 5.19 (d, *J* = 3.5 Hz, 1H, H-1''), 5.03 - 4.99 (m, 4H, -C*H*$_2$Bn), 4.90 (appt, *J* = 4.4 Hz, 1H, H-2'), 4.52 (appt, *J* = 4.8 Hz, 1H, H-3'), 4.39 (ddd, *J* = 10.7, 6.0, 4.2 Hz, 1H, H-5'a), 4.31 – 4.23 (m, 1H, H-4'), 4.21 – 4.08 (m, 2H, H-5'b + H-4''), 4.00 (appt, J = 6.0 Hz, 1H, H-3''), 3.92 (dd, *J* = 5.0, 3.5 Hz, 1H, H-2''), 3.78 (d, *J* = 12.1 Hz, 1H, H-5''a), 3.60-3.54 (m, 1H, H-5''b), 0.90 (s, 9H, -tBu), 0.90 (s, 9H, -tBu), 0.85 (s, 9H, -tBu), 0.11

(s, 3H, -CH_3), 0.08 (s, 3H, -CH_3), 0.07 (s, 3H, -CH_3), 0.06 (s, 3H, -CH_3), -0.00 (d, 3H, -CH_3), -0.02 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 165.0, 152.8, 151.5, 149.6, 142.6, 135.7, 135.6, 133.2, 129.1, 128.9, 128.8, 128.2, 128.1, 123.7, 103.5, 88.1, 83.5, 82.9, 80.2, 77.5, 73.8, 70.8, 69.8, 65.9, 61.4, 26.2, 26.2, 26.0, 18.6, 18.3, 18.3, -4.0, -4.2, -4.3, -4.5, -4.5, -4.8.

^{31}P NMR (202 MHz, $CDCl_3$) δ 0.06

HRMS (ESI) m/*z* calcd for $C_{54}H_{81}N_5O_{12}Si_3P$ ($[M+H]^+$) 1106.4927, found 1106.4918.



Compound 2-29: To a -78°C solution of compound **2-27** (250 mg, 0.23 mmol, 1 eq) in THF (2 mL) was added diisopropylethylamine (0.047 mL, 0.27 mmol, 1.2 eq) followed by 2-Cyanoethyl N,N-diisopropylchlorophosphoramidite (**2-28**) (64 mg, 0.27 mmol, 1.2 eq). The -78°C bath was removed and the reaction was allowed to warm to room temperature and stirred for 3 h. The reaction mixture was filtered through celite, concentrated. The product was purified by silica column chromatography (0.67 in 50:50 hexanes:ethyl acetate with 1% triethylamine) to give compound **2-29** (310 mg, 99%) as a mixture of 2 diastereomers at the phosphoramidite.

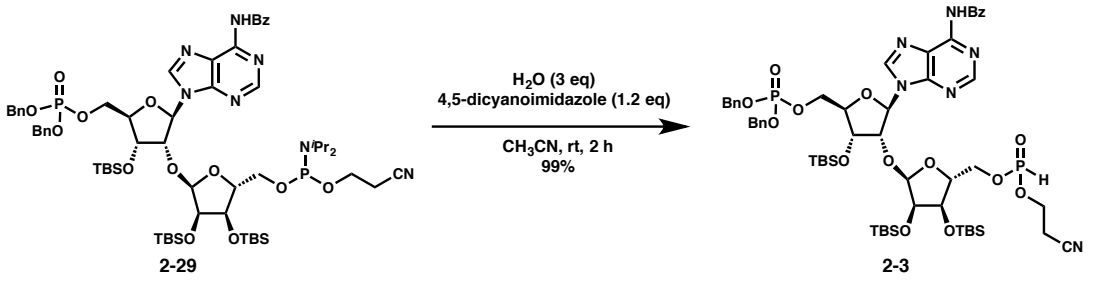
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.91 (s, 2H, -N*H*Bz), 8.75 (s, 1H, H-8), 8.75 (s, 1H, H-8), 8.27 (s, 1H, H-2), 8.24 (s, 1H, H-2), 8.03 – 7.98 (m, 4H, -Bz), 7.66 – 7.58 (m, 2H, -Bz), 7.57 – 7.50 (m, 4H, -Bz), 7.31-7.28 (m, 20H, -Bn), 6.16 (d, *J* = 3.0 Hz, 1H, H-1'), 6.15 (d, *J* = 3.1 Hz, 1H, H-1'), 5.22 (d, *J* = 3.9 Hz, 1H, H-1''), 5.21 (d, *J* = 3.8 Hz, 1H, H-1''), 5.02 – 4.94 (m, 8H, -C*H*$_2$Bn), 4.81 – 4.74 (m, 2H, H-2'), 4.55 – 4.48 (m, 2H, H-3'), 4.43 – 4.34 (m, 2H, H-5'a), 4.32

– 4.25 (m, 2H, H-4'), 4.22 – 4.14 (m, 4H, H-5'b + H-4''), 4.07 – 3.99 (m, 2H, H-3''), 3.99 – 3.91 (m, 2H, H-2''), 3.87 - 3.52 (m, 12H, H-5''a+b + -C*H*$_2$CH$_2$CN + -C*H*(CH$_3$)$_2$), 2.59 (appq, *J* = 6.1 Hz, 4H, -CH$_2$C*H*$_2$CN), 1.20 – 1.12 (m, 24H, -CH(C*H*$_3$)$_2$), 0.90 (s, 9H, -tBu), 0.90 (s, 9H, -tBu), 0.89 (s, 9H, -tBu), 0.89 (s, 9H, -tBu), 0.83 (s, 9H, -tBu), 0.83 (s, 9H, -tBu), 0.12 (s, 3H, -CH$_3$), 0.12 (s, 3H, -CH$_3$), 0.08 (s, 3H, -CH$_3$), 0.08 (s, 3H, -CH$_3$), 0.06 (s, 3H, -CH$_3$), 0.06 (s, 3H, -CH$_3$), 0.06 (s, 3H, -CH$_3$), 0.06 (s, 3H, -CH$_3$) 0.01 (s, 3H, -CH$_3$), 0.00 (s, 3H, -CH$_3$), -0.01 (s, 3H, -CH$_3$), -0.01 (s, 3H, -CH$_3$).

^{13}C NMR (125 MHz, CDCl$_3$) δ 164.5, 152.8, 151.3, 151.2, 149.6, 149.6, 142.4, 142.3, 135.8, 135.8, 133.9, 133.0, 133.0, 129.1, 128.8, 128.2, 128.2, 128.0, 123.8, 123.7, 117.8, 117.7, 103.6, 103.5, 88.7, 88.7, 84.6, 84.5, 84.2, 84.1, 80.5, 80.4, 73.8, 71.3, 71.1, 70.5, 70.4, 69.7, 69.6, 69.6, 69.6, 65.6, 65.6, 63.2, 63.2, 63.1, 63.0, 58.7, 58.6, 58.5, 58.4, 43.4, 43.3, 43.3, 43.2, 26.2, 26.2, 26.2, 26.0, 26.0, 25.0, 24.9, 24.9, 24.8, 24.8, 20.6, 20.6, 20.6, 20.6, 18.6, 18.5, 18.3, -4.1, -4.2, -4.2, -4.2, -4.3, -4.3, -4.4, -4.4, -4.5, -4.8, -4.8.

^{31}P NMR (202 MHz, CDCl$_3$) δ 149.67, 0.15, 0.12

Compound **2-29** hydrolyzed when subjected to ESI-MS.



Compound 2-3: To a 7 mL reaction vial was added compound **2-29** (170 mg, 0.13 mmol, 1 eq) and CH$_3$CN (1 mL). Next, 4,5-dicyanoimidazole (18.4 mg, 0.16 mmol, 1.2 eq) and deionized water (7.0 mg, 0.39 mmol, 3 eq) were added and the reaction was stirred at room temperature for 2 h. The reaction mixture was poured into saturated aqueous NaHCO$_3$, extracted three times with

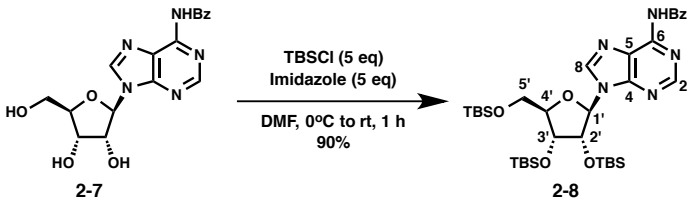
EtOAc, dried through Na_2SO_4, and concentrated to give compound **2-3** (158 mg, 99%) as a brown oil and a mixture of diastereomers at the H-phosphonate. The compound was of sufficient purity to be used in the next step without further purification.

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.91 (s, 1H, -N*H*Bz), 8.90 (s, 1H, N*H*Bz), 8.78 (s, 1H, H-8), 8.77 (s, 1H, H-8), 8.29 (s, 1H, H-2), 8.26 (s, 1H, H-2), 8.01 (d, *J* = 6.7 Hz, 4H, -Bz), 7.67-7.60 (m, 2H, -Bz) 7.57 – 7.50 (m, 4H, -Bz), 7.33-7.29 (m, 20H, -Bn), 6.93 (d, J = 729 Hz, 1H, -PH), 6.93 (d, J = 707 Hz, 1H, -PH), 6.16 (d, *J* = 3.8 Hz, 2H, H-1'), 5.28 (d, *J* = 3.2 Hz, 1H, H-1''), 5.25 (d, *J* = 3.3 Hz, 1H, H-1''), 5.04-4.97 (m, 8H, -Bn), 4.88 (appt, *J* = 4.2 Hz, 1H, H-2'), 4.84 (appt, *J* = 4.2 Hz, 1H, H-2'), 4.56 – 4.45 (m, 2H, H-3'), 4.42 - 3.93 (m, 20H), 2.77 – 2.67 (m, 4H), 0.90 (s, 9H), 0.90 (s, 9H), 0.90 (s, 9H), 0.90 (s, 9H), 0.86 (s, 9H), 0.86 (s, 9H), 0.12 (s, 3H), 0.11 (s, 3H), 0.08 (s, 3H), 0.08 (s, 3H), 0.08 (s, 3H), 0.08 (s, 3H), 0.08 (s, 3H), 0.07 (s, 3H), 0.02 (s, 3H), 0.02 (s, 3H), -0.01 (s, 3H), -0.01 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 164.6, 164.6, 152.7, 152.6, 151.3, 151.3, 149.7, 149.6, 142.2, 142.1, 135.7, 135.6, 133.8, 132.9, 129.0, 128.7, 128.7, 128.1, 128.1, 128.0, 127.9, 123.8, 103.4, 103.2, 88.0, 87.9, 82.8, 82.7, 80.5, 80.3, 73.3, 70.8, 70.6, 70.5, 69.6, 65.6, 64.9, 64.8, 60.1, 59.9, 47.0, 46.2, 22.7, 22.5, 20.1, 20.0, 19.2, 18.5, 18.5, 18.2, 18.2, 18.2, -4.0, -4.1, -4.3, -4.3, -4.4, -4.4, -4.5, -4.6, -4.6, -4.7, -4.9, -4.9.

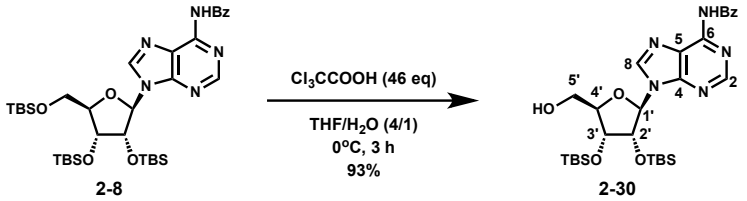
^{31}P NMR (202 MHz, $CDCl_3$) δ 9.40, 9.28, -0.17, -0.18

HRMS (ESI) m/*z* calcd for $C_{57}H_{85}N_6O_{14}Si_3P_2$ ($[M+H]^+$) 1223.4907, found 1223.4899.



2-7 2-8

Compound 2-8: A modified version of the procedure of Aritomo *et al.*[23] was followed. To a 20 mL scintillation vial was added N-Benzoyladenosine (**2-7**) (500 mg, 1.35 mmol, 1 eq) and imidazole (458 mg, 6.73 mmol, 5 eq). Next, DMF (5 mL) was added followed by TBSCl (1.01 g, 6.73 mmol, 5 eq) and the reaction was stirred at room temperature for 14 h. The reaction was concentrated and purified by silica column chromatography (R_f 0.32 in 70:30 hexanes:ethyl acetate) to yield compound **2-8** as a yellow oil (866 mg, 90%).



Compound 2-30: A modified version of the procedure of Zhu *et al.*[24] was followed. To a 250 mL round bottom flask was added compound **2-8** (3.0 g, 4.2 mmol, 1 eq) and THF (60 mL). The solution was cooled to 0 ^oC. Next, a 0 ^oC solution of trichloroacetic acid (32 g, 193 mmol, 46 eq) in water (15 mL) was added slowly dropwise by cannula. The reaction was stirred at 0 ^oC for 3 h and then quenched by dropwise addition of a 0 ^oC solution of saturated sodium bicarbonate. The reaction was extracted three times with CH_2Cl_2, dried through Na_2SO_4, concentrated and purified by silica column chromatography (R_f 0.10 in 50:50 hexanes:ethyl acetate) to yield compound **2-30** as a white solid (2.02 g, 93%).

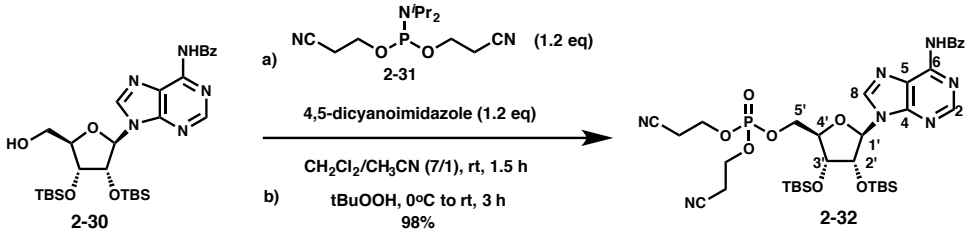
^{1}H NMR (500 MHz, $CDCl_3$) δ 9.04 (s, 1H, -N*H*Bz), 8.83 (s, 1H, H-8), 8.06 – 8.01 (m, 3H, H-2 + -Bz), 7.63 (t, *J* = 7.4 Hz, 1H, -Bz), 7.54 (t, *J* = 7.6 Hz, 2H, -Bz), 6.13 (dd, *J* = 12.0, 2.0 Hz, 1H, -OH), 5.86 (d, *J* = 7.9 Hz, 1H, H-1'), 5.06 (dd, *J* = 7.9, 4.5 Hz, 1H, H-2'), 4.35 (d, *J* = 4.5 Hz, 1H, H-3'), 4.21-4.18 (m, 1H, H-4'), 3.98 (appdt, *J* = 13.1, 2.0 Hz, 1H, H-5'a), 3.74 (ddd, *J* = 13.3, 12.1, 1.6 Hz, 1H, H-5'b), 0.96 (s, 9H, -tBu), 0.75 (s, 9H, -tBu), 0.14 (s, 3H, -CH_3), 0.13 (s, 3H, -CH_3), -0.13 (s, 3H, -CH_3), -0.64 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 164.4, 152.4, 150.6, 150.5, 143.4, 133.7, 133.1, 129.1, 128.0, 124.5, 91.4, 89.8, 74.2, 74.0, 63.1, 25.9, 25.8, 18.2, 17.9, -4.4, -4.4, -4.5, -5.7.

HRMS (ESI) m/*z* calcd for $C_{29}H_{46}N_5O_5Si_2$ ($[M+H]^+$) 600.3038, found 600.3030.

IR (neat) ν 3268, 2929, 2850, 1702, 1611, 1510, 1459, 1252, 1157, 1098, 1063, 836, 777 cm^{-1}.

$[a]_D^{23}$ -46.1 (*c* = 1.1, $CHCl_3$)



Compound 2-32: To a 25 mL round-bottomed flask was added compound **2-30** (500 mg, 0.83 mmol, 1 eq) and phosphoramidite **2-31** (301 mg, 1.0 mmol, 1.2 eq). The compounds were co-evaporated three times on a N_2-filled rotary evaporator with dry CH_3CN and then dried over P_2O_5 overnight. Separately, 4,5-dicyanoimidazole (118 mg, 1.0 mmol, 1.2 eq) was evaporated and dried in a similar manner. To compounds **2-30** and **2-31** was added dry CH_2Cl_2 (7 mL), followed by a solution of the 4,5-dicyanoimidazole in CH_3CN (1 mL). After 1.5 h, the reaction was cooled to 0 ^oC and a solution of t-butyl peroxide (1.0 mL of a 5.5 M solution in decane, 5.5 mmol, 5.5 eq) was added. The reaction was warmed to room temperature and stirred for 3 h. Next, the reaction was poured into a saturated aqueous $NaHCO_3$ solution, extracted three times with EtOAc, dried through Na_2SO_4 and concentrated. The product was purified by silica gel chromatography (R_f=0.11 in 100% EtOAc) to give compound **2-32** as a white foam (640 mg, 98%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.90 (s, 1H, -N*H*Bz), 8.82 (s, 1H, H-8), 8.25 (s, 1H, H-2), 8.02 (d, *J* = 8.2 Hz, 2H, -Bz), 7.63 (t, *J* = 7.4 Hz, 1H, -Bz), 7.54 (t, *J* = 7.6 Hz, 2H, -Bz), 6.00 (d, *J* =

4.0 Hz, 1H, H-1'), 4.92 (appt, *J* = 4.1 Hz, 1H, H-2'), 4.51 (dd, *J* = 10.6, 4.3 Hz, 1H, H-5'a), 4.43 – 4.35 (m, 2H, H-3' + H-5'), 4.35 – 4.21 (m, 5H, H-4' + -C*H*$_2$CH$_2$CN), 2.81 – 2.73 (m, 4H, -CH$_2$C*H*$_2$CN), 0.94 (s, 9H, -tBu), 0.85 (s, 9H, -tBu), 0.14 (s, 3H, -CH$_3$), 0.12 (s, 3H, -CH$_3$), 0.03 (s, 3H, -CH$_3$), -0.12 (s, 3H, -CH$_3$).

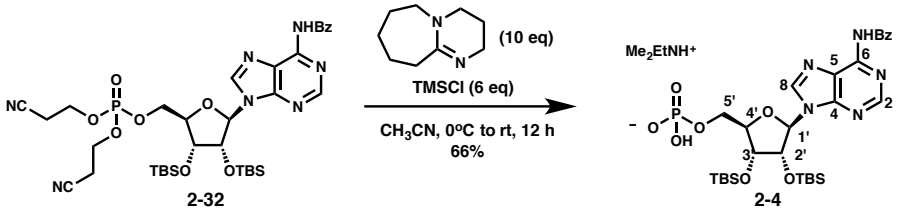
^{13}C NMR (125 MHz, CDCl$_3$) δ 164.8, 152.9, 151.7, 149.9, 142.5, 133.8, 133.1, 129.1, 128.1, 116.5, 90.1, 82.6 (d, J = 8.5 Hz), 74.8, 71.6, 67.0 (d, J = 5.0 Hz), 62.7 (t, J = 4.8 Hz), 26.0, 26.0, 19.9 (d, J = 7.3 Hz), 18.3, 18.1, -4.1, -4.4, -4.6, -4.6.

^{31}P NMR (202 MHz, CDCl$_3$) δ -0.97

HRMS (ESI) m/z calcd for $C_{35}H_{53}N_7O_8PSi_2$ ([M+H]$^+$) 786.3232, found 786.3225

IR (neat) ν 2954, 2919, 2857, 1701, 1611, 1576, 1458, 1250, 1036, 997, 884, 782 cm^{-1}

[a]$_D^{23}$ -6.7 (*c* = 1.05, CHCl$_3$)



Compound 2-4: To a 25 mL round-bottomed flask was added compound **2-32** (500 mg, 0.64 mmol, 1 eq), followed by CH$_3$CN (7 mL). The solution was cooled to 0 ^{o}C, and 1,8-Diazabicyclo[5.4.0]undec-7-ene (974 mg, 6.4 mmol, 10 eq) was added dropwise. Finally, TMSCl (413 mg, 3.8 mmol, 6 eq) was added. After 5 min, the reaction was removed from the 0 ^{o}C bath, allowed to warm to room temperature, and stirred for 12 h. The reaction was concentrated and applied to a Dowex 50W-X8 cation exchange resin (Me$_2$EtNH$^+$ form) and eluted with MeOH. The product was purified by C-18 chromatography to yield compound **2-4** as the dimethyl ethyl ammonium salt (316 mg, 66%) as a white foam.

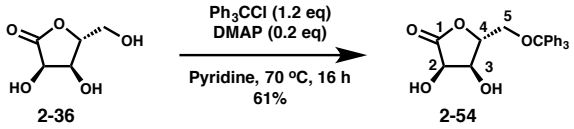
^{1}H NMR (500 MHz, CD_3OD) δ 8.90 (s, 1H, H-8), 8.75 (s, 1H, H-2), 8.11 (d, J=7.0, 2H, -Bz), 7.67 (t, J=7.5 Hz, 1H, -Bz), 7.58 (t, J=7.5 Hz, 2H, -Bz), 6.26 (d, *J* = 6.8 Hz, 1H, H-1'), 4.99 – 4.85 (m, 1H, H-2'), 4.49 (dd, *J* = 4.4, 1.8 Hz, 1H, H-3'), 4.32 – 4.24 (m, 1H, H-4'), 4.24 – 4.18 (m, 1H, H-5'a), 4.13 (ddd, *J* = 11.3, 5.0, 3.2 Hz, 1H, H-5'b), 3.16 (q, *J* = 7.3 Hz, 2H, $CH_3CH_2N(CH_3)_2$), 2.85 (s, 6H, $CH_3CH_2N(CH_3)_2$), 1.33 (t, J = 7.3 Hz, 3H, $CH_3CH_2N(CH_3)_2$), 1.01 (s, 9H, -tBu), 0.77 (s, 9H, -tBu), 0.22 (s, 3H, $-CH_3$), 0.19 (s, 3H, $-CH_3$), 0.03 (s, 3H, $-CH_3$), -0.31 (s, 3H, $-CH_3$).

^{13}C NMR (125 MHz, CD_3OD) δ 168.1, 153.8, 153.3, 151.2, 144.7, 135.1, 133.8, 129.8, 129.4, 124.8, 88.8, 87.5 (d, J=9.0 Hz), 77.6, 74.7, 65.7 (d, J=5.8 Hz), 53.9, 42.8, 26.4, 26.2, 18.9, 18.7, 10.0, -4.2, -4.3, -4.3, -5.1.

^{31}P NMR (202 MHz, CD_3OD) δ 1.93

HRMS (ESI) m/z calcd for $C_{29}H_{47}N_5O_8PSi_2$ ($[M+H]^+$) 680.2701, found 680.2706

$[a]_D^{22}$ -46.1 (*c* = 1.2, EtOH)



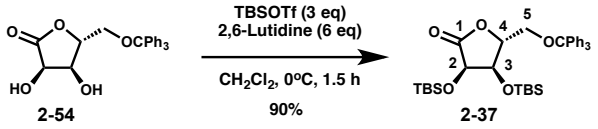
2-36 2-54

Compound 2-54: A modified version of the procedure of Ireland *et al.*[25] was followed. To a 20 mL reaction vial was added compound D-ribonolactone (**2-36**) (500 mg, 3.38 mmol, 1 eq) and pyridine (10 mL). Next, triphenylmethyl chloride (1.13 g, 4.05 mmol, 1.2 eq) was added followed by and 4-dimethylaminopyridine (82 mg, 0.675 mmol, 0.2 eq). The reaction was stirred at 70°C for 16 h at which point it was concentrated and purified by silica column chromatography to yield compound **2-54** (800 mg, 61%) as a white solid.

^{1}H NMR (500 MHz, Acetone-*d_6*) δ 7.46 -7.26 (m, 15H, -CPh_3), 4.76 (d, *J* = 5.4 Hz, 1H, H-2), 4.47 (apptd, *J* = 3.4, 1.0 Hz, 1H, H-4), 4.27 (d, *J* = 5.3 Hz, 1H, H-3), 3.55 (dd, *J* = 10.9, 3.4 Hz, 1H, H-5a), 3.28 (dd, *J* = 10.9, 3.3 Hz, 1H, H-5b).

^{13}C NMR (125 MHz, Acetone-d_6) δ 176.2, 144.4, 129.4, 128.8, 128.2, 88.2, 84.3, 70.9, 69.8, 64.1.

HRMS (ESI) m/z calcd for $C_{24}H_{22}O_5Na$ ($[M+Na]^+$) 413.1365, found 413.1361

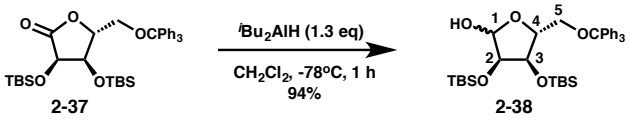


Compound 2-37: To a 20 mL reaction vial was added compound **2-54** (370 mg, 0.95 mmol, 1 eq) and CH_2Cl_2 (10 mL). The vial was cooled to 0°C and 2,6-Lutidine (0.66 mL, 5.7 mmol, 6 eq) was added. Finally, TBSOTf (0.65 mL, 2.9 mmol, 3 eq) was added over 30 min. The reaction was stirred for 1 h at 0°C at which point a saturated aqueous $NaHCO_3$ solution. The mixture was extracted three times with $CHCl_3$, dried through Na_2SO_4, concentrated and purified by silica column chromatography to yield known[26,29] compound **2-37** (527 mg, 90%) as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.47 – 7.40 (m, 6H, -CPh_3), 7.38 – 7.31 (m, 6H, -CPh_3), 7.31 – 7.25 (m, 3H, -CPh_3), 4.72 (d, *J* = 5.2 Hz, 1H, H-2), 4.33 (ddd, *J* = 3.6, 2.7, 1.1 Hz, 1H, H-4), 3.99 (dd, *J* = 5.1 Hz, 1.1 Hz, 1H, H-3), 3.66 (dd, *J* = 11.0, 3.6 Hz, 1H, H-5a), 3.25 (dd, *J* = 11.0, 2.8 Hz, 1H, H-5b), 0.97 (s, 9H, -tBu), 0.84 (s, 9H, -tBu), 0.23 (s, 3H, -CH_3), 0.15 (s, 3H, -CH_3), 0.05 (s, 3H, -CH_3), -0.02 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 175.2, 143.2, 128.6, 128.2, 127.5, 87.7, 84.7, 72.2, 70.4, 62.4, 26.0, 25.8, 18.5, 18.2, -4.5, -4.5, -4.7, -5.0.

HRMS (ESI) m/z calcd for $C_{36}H_{50}NaO_5Si_2$ ($[M+Na]^+$) 641.3089, found 641.3064.



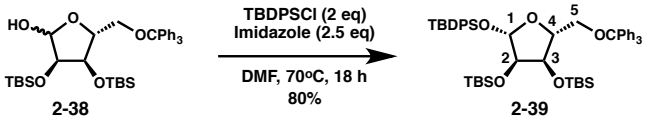
Compound 2-38: To a -78°C solution of compound **2-37** (200 mg, 0.32 mmol, 1 eq) in CH_2Cl_2 (3.5 mL) was added diisobutylaluminum hydride (0.42 mL of 1 M solution in hexanes, 0.42 mmol, 1.3 eq). The reaction was stirred at -78°C for 1 h, at which point the reaction was slowly quenched with MeOH. The reaction mixture was poured into a saturated aqueous solution of potassium sodium tartrate, extracted three times with EtOAc and CH_2Cl_2, dried through Na_2SO_4, concentrated and purified by silica column chromatography to give compound **2-38** as a white solid (187 mg, 94%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.52 – 7.41 (m, 6H major and minor, -CPh_3), 7.38 – 7.29 (m, 6H, major and minor, -CPh_3), 7.29 – 7.21 (m, 3H major and minor, -CPh_3), 5.22 (dd, *J* = 4.4, 1.1 Hz, 1H minor, H-1b), 5.16 (dd, *J* = 11.3, 4.2 Hz, 1H major, H-1a), 4.42 (dd, *J* = 7.3, 4.0 Hz, 1H, H-3b minor), 4.26 (d, *J* = 11.3 Hz, 1H major, -OHa), 4.23 (appt, *J*= 4.0 Hz, 1H major, H-4a), 4.19 (appt, *J* = 4.5 Hz, 1H major, H-2a), 4.10 (ddd, *J* = 6.9, 3.8, 2.8 Hz, 1H minor, H-4b), 3.99 (d, *J* = 4.0 Hz, 1H minor, H-2b), 3.95 (d, *J* = 4.6 Hz, 1H major, H-3a), 3.51 (dd, *J* = 10.3, 2.7 Hz, 1H minor, H-5b), 3.30 (dd, *J* = 10.3, 5.0 Hz, 1H major, H-5a), 3.14 (dd, *J* = 10.3, 3.3 Hz, 1H major, H-5a'), 3.14-3.11 (m, 1H, minor, H-5b') 2.79 (d, *J* = 4.6 Hz, 1H minor, -OHb), 0.96 (s, 9H major, -tBu), 0.94 (s, 9H minor, -tBu), 0.88 (s, 9H major, -tBu), 0.78 (s, 9H minor, -tBu), 0.16 (s, 3H major, -CH_3), 0.14 (s, 3H minor, -CH_3), 0.13 (s, 3H major, -CH_3), 0.12 (s, 3H minor, -CH_3), 0.07 (s, 3H major, -CH_3), 0.01 (s, 3H major, -CH_3), 0.00 (s, 3H minor, -CH_3), -0.16 (s, 3H minor, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 143.8 (major), 143.8 (minor), 129.0 (minor), 128.8 (major), 128.0 (major), 127.9 (minor), 127.3 (major), 127.2 (minor), 102.2 (minor), 97.9 (major), 87.1

(major), 87.0 (minor), 84.4 (major), 81.8 (minor), 74.6 (major and minor), 72.6 (major), 71.8 (major), 63.8 (major), 63.4 (minor), 26.0 (major), 26.0 (minor), 25.9 (minor), 25.9 (major), 18.4 (major), 18.3 (minor), 18.1 (major and minor), -4.1 (minor), -4.3 (minor), -4.4 (minor), -4.5 (major), -4.6 (major), -4.7 (major), -4.9 (major), -4.9 (minor).

HRMS (ESI) m/z calcd for $C_{36}H_{52}O_5Si_2Na$ ($[M+Na]^+$) 643.3251, found 643.3259.

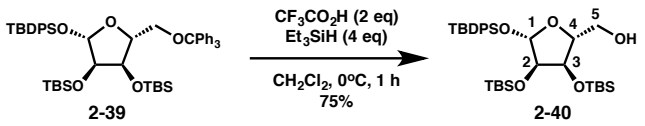


Compound 2-39: To a 7 mL reaction vial was added compound **2-38** (100 mg, 0.16 mmol, 1 eq) and DMF (1 mL). Imidazole (27 mg, 0.4 mmol, 2.5 eq) was added, followed by tert-Butyl(chloro)diphenylsilane (0.083 mL, 0.32 mmol, 2 eq) and the reaction was heated to 70°C for 18 h. The reaction was quenched with a saturated aqueous NH_4Cl solution, extracted three times with CH_2Cl_2, dried through Na_2SO_4, concentrated and purified by silica column chromatography to yield compound **2-39** as a white foam (110 mg, 80%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.86 – 7.78 (m, 2H, -TBDPS), 7.73 – 7.66 (m, 2H, -TBDPS), 7.63 – 7.53 (m, 6H, -CPh_3), 7.49 – 7.36 (m, 6H, -TBDPS), 7.30 – 7.24 (m, 6H, -CPh_3), 7.24 – 7.17 (m, 3H, -CPh_3), 5.17 (s, 1H, H-1), 4.40 (dd, *J* = 8.5, 3.6 Hz, 1H, H-3), 4.18 (ddd, *J* = 8.4, 6.2, 2.0 Hz, 1H, H-4), 3.79 (d, *J* = 3.6 Hz, 1H, H-2), 3.44 (dd, *J* = 10.1, 2.0 Hz, 1H, H-5a), 3.15 (dd, *J* = 10.2, 6.2 Hz, 1H, H-5b), 1.03 (s, 9H, -tBu), 0.77 (s, 9H, -tBu), 0.76 (s, 9H, -tBu), -0.02 (s, 3H, -CH_3), -0.12 (s, 3H, -CH_3), -0.14 (s, 3H, -CH_3), -0.39 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 144.3, 135.9, 135.7, 134.1, 132.9, 130.0, 130.0, 129.1, 128.0, 128.0, 127.8, 126.9, 102.1, 86.6, 80.6, 77.8, 72.1, 65.3, 27.0, 26.0, 25.9, 19.3, 18.3, 18.1, -4.0, -4.7, -4.9, -5.0.

HRMS (ESI) m/z calcd for $C_{52}H_{70}O_5Si_3Na$ ($[M+Na]^+$) 881.4429, found 881.4441.

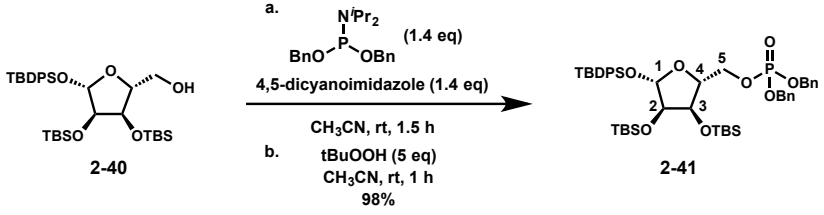


Compound 2-40: To a 7 mL reaction vial was added compound **2-39** (90 mg, 0.105 mmol, 1 eq) and CH_2Cl_2 (1 mL). The solution was cooled to 0°C and Et_3SiH (0.067 mL, 0.419 mmol, 4 eq) was added. Finally, CF_3CO_2H (0.016 mL, 0.209 mmol, 2 eq) was added in a solution of CH_2Cl_2 (0.1 mL) and the reaction was stirred at 0°C for 1 h. The reaction was quenched with a solution of saturated aqueous $NaHCO_3$, extracted three times with CH_2Cl_2, dried through Na_2SO_4, concentrated, and purified by silica column chromatography to give compound **2-40** (44 mg, 75%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.72 – 7.69 (m, 2H, -TBDPS), 7.66 – 7.62 (m, 2H, -TBDPS), 7.48 – 7.36 (m, 6H, -TBDPS), 5.05 (s, 1H, H-1), 4.60 (dd, *J* = 8.3, 3.6 Hz, 1H, H-3), 4.04 (ddd, *J* = 8.3, 2.6, 2.5 Hz, 1H, H-4), 3.86 (dd, *J* = 12.0, 2.5 Hz, 1H, H-5a), 3.78 (d, *J* = 3.6 Hz, 1H, H-2), 3.60 (dd, *J* = 12.2, 2.6 Hz, 1H, H-5b), 2.01 (bs, 1H, -OH), 1.07 (s, 9H, -tBu), 0.92 (s, 9H, -tBu), 0.76 (s, 9H, -tBu), 0.15 (s, 3H, $-CH_3$), 0.11 (s, 3H, $-CH_3$), -0.10 (s, 3H, $-CH_3$), -0.34 (s, 3H, $-CH_3$).

^{13}C NMR (125 MHz, $CDCl_3$) δ 135.8, 133.3, 132.7, 130.3, 130.2, 128.1, 101.8, 82.0, 78.2, 70.3, 60.8, 26.9, 26.1, 25.8, 19.2, 18.3, 18.1, -4.1, -4.7, -4.8, -4.8.

HRMS (ESI) m/z calcd for $C_{33}H_{56}O_5NaSi_3$ ($[M+Na]^+$) 639.3333, found 639.3342.



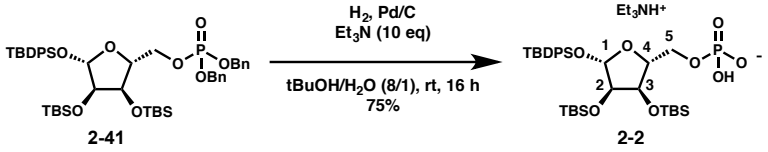
Compound 2-41: A 7 mL vial containing compound **2-40** (35 mg, 0.057 mmol, 1 eq) was azeotroped twice in a nitrogen filled rotary evaporator with dry CH_3CN and dried on a high vacuum for 20 min. To compound **2-40** was added dibenzyl N,N-diisopropylphosphoramidite (28 mg, 0.080 mmol, 1.4 eq) in a solution of CH_3CN (0.4 mL). Next, dicyanoimidazole (9.4 mg, 0.080 mmol, 1.4 eq) was added in a solution of CH_3CN (0.1 mL), the reaction was stirred at room temperature for 1.5 h at which point tBuOOH (0.052 mL of a 5.5 M solution in decane, 0.285 mmol, 5 eq) was added. The reaction stirred an additional 1 h at room temperature at which point it was quenched with saturated aqueous $NaHCO_3$, extracted three times with CH_2Cl_2, dried through Na_2SO_4, and concentrated. The reaction was purified by silica column chromatography to give compound **2-41** (49 mg, 98%).

^{1}H NMR (500 MHz, $CDCl_3$) d 7.76 – 7.69 (m, 2H, -TBDPS), 7.66 – 7.60 (m, 2H, -TBDPS), 7.47 – 7.28 (m, 16H, -TBDPS + -Bn), 5.10 – 4.96 (m, 5H, H-1 + CH_2Ph), 4.40 (dd, *J* = 8.0, 3.6 Hz, 1H, H-3), 4.35 – 4.29 (m, 1H, H-5a), 4.20 – 4.09 (m, 2H, H-4 + H-5b), 3.83 (d, *J* = 3.6 Hz, 1H, H-2), 1.05 (s, 9H, -tBu), 0.91 (s, 9H, -tBu), 0.77 (s, 9H, -tBu), 0.12 (s, 3H, -CH_3), 0.10 (s, 3H, -CH_3), -0.08 (s, 3H, -CH_3), -0.32 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) d 135.8, 135.7, 133.6, 132.9, 130.1, 129.9, 128.6, 128.5, 128.0, 128.0, 128.0, 127.9, 101.9, 79.6 (d, *J* = 8.5 Hz), 77.8, 72.0, 69.3 (d, *J* = 5.6 Hz), 69.2 (d, *J* = 5.6 Hz), 68.6 (d, *J* = 5.6 Hz), 26.9, 26.0, 25.8, 19.2, 18.2, 18.0, -4.1, -4.6, -4.8.

^{31}P NMR (202 MHz, $CDCl_3$) d -0.29

HRMS (ESI) m/z calcd for $C_{47}H_{69}O_8PSi_3Na$ ($[M+Na]^+$) 899.3936, found 899.3935.



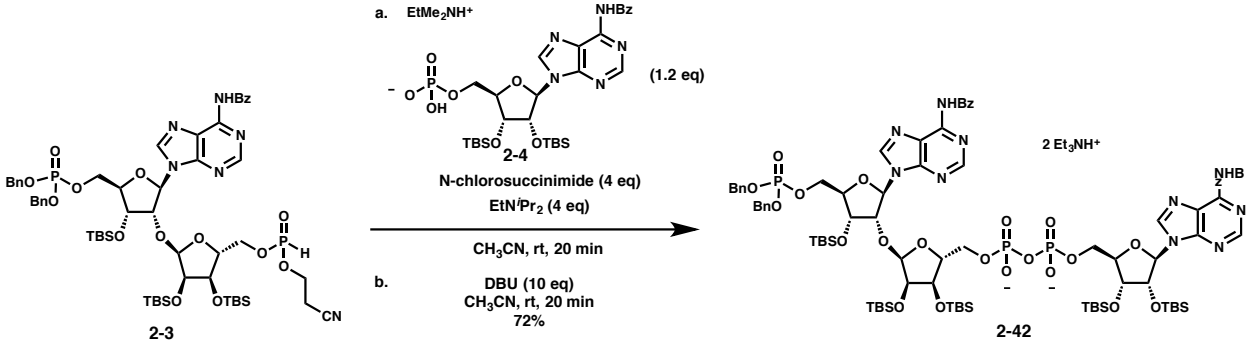
Compound 2-2: To a 20 mL vial containing compound **2-41** (48 mg, 0.055 mmol, 1 eq) was added 10% Palladium on activated charcoal (50 mg). Next, tBuOH (4 mL) and deionized water (0.5 mL) were added, followed by Et_3N (0.075 mL, 0.55 mmol, 10 eq). The reaction was fitted with a ballon of hydrogen gas, purged for 1 min and stirred at room temperature for 16 h. After 16 h, the reaction was filtered through celite, concentrated, and purified by reverse phase chromatography to give compound **2-2** as a pale yellow oil (32 mg, 75%).

^{1}H NMR (500 MHz, CD_3OD) δ 7.79 – 7.71 (m, 2H, -TBDPS), 7.68 – 7.61 (m, 2H, -TBDPS), 7.52 – 7.35 (m, 6H, -TBDPS), 5.03 (s, 1H, H-1), 4.34 (dd, *J* = 7.9, 3.6 Hz, 1H, H-3), 4.24 (ddd, *J* = 10.5, 6.1, 2.7 Hz, 1H, H-5a), 4.12 (apptd, *J* = 8.1, 2.7 Hz, 1H, H-4), 3.99 (ddd, *J* = 10.5, 8.3, 5.0 Hz, 1H, H-5b), 3.89 (d, *J* = 3.6 Hz, 1H, H-2), 3.15 (q, *J* = 7.3 Hz, 6H, -(CH_3CH_2)$_3$N), 1.28 (t, *J* = 7.3 Hz, 9H, -(CH_3CH_2)$_3$N), 1.06 (s, 9H, -tBu), 0.94 (s, 9H, -tBu), 0.78 (s, 9H, -tBu), 0.18 (s, 3H, -CH_3), 0.15 (s, 3H, -CH_3), -0.06 (s, 3H, -CH_3), -0.28 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, CD_3OD) δ 137.0, 136.8, 134.5, 133.9, 131.2, 131.0, 129.0, 128.9, 103.0, 82.4 (d, *J* = 8.8 Hz), 79.08, 74.34, 69.2 (d, *J* = 5.4 Hz), 47.4, 27.3, 26.6, 26.2, 19.9, 19.0, 18.8, 9.1, -4.0, -4.3, -4.4, -4.5.

^{31}P NMR (202 MHz, CD_3OD) δ 2.11

HRMS (ESI) m/z calcd for $C_{33}H_{58}O_8Si_3P$ ($[M+H]^+$) 697.3177, found 697.3192.



Compound 2-42: To a 7 mL reaction vial was added compound **2-3** (68 mg, 0.056 mmol, 1 eq) and compound **2-4** (50 mg, 0.067 mmol, 1.2 eq). The mixture was co-evaporated three times with dry CH_3CN and dried over P_2O_5 overnight. Compounds **2-3** and **2-4** were dissolved in dry CH_3CN (0.5 mL) containing N,N-diisopropylethylamine (29 mg, 0.22 mmol, 4 eq) and N-chlorosuccinimide (30 mg, 0.22 mmol, 4 eq) was added and the reaction was stirred for 20 min at room temperature. Diazabicyclo[5.4.0]undec-7-ene (84 mg, 0.556 mmol, 10 eq) was added and the reaction was stirred for 20 min at room temperature. The solution was concentrated and purified by reverse phase chromatography. The purified compound was passed through a short column of Dowex 50W-X8 (Et_3NH^+ form) to yield compound **2-42** as the triethylammonium salt as a pale yellow foam (82 mg, 72%).

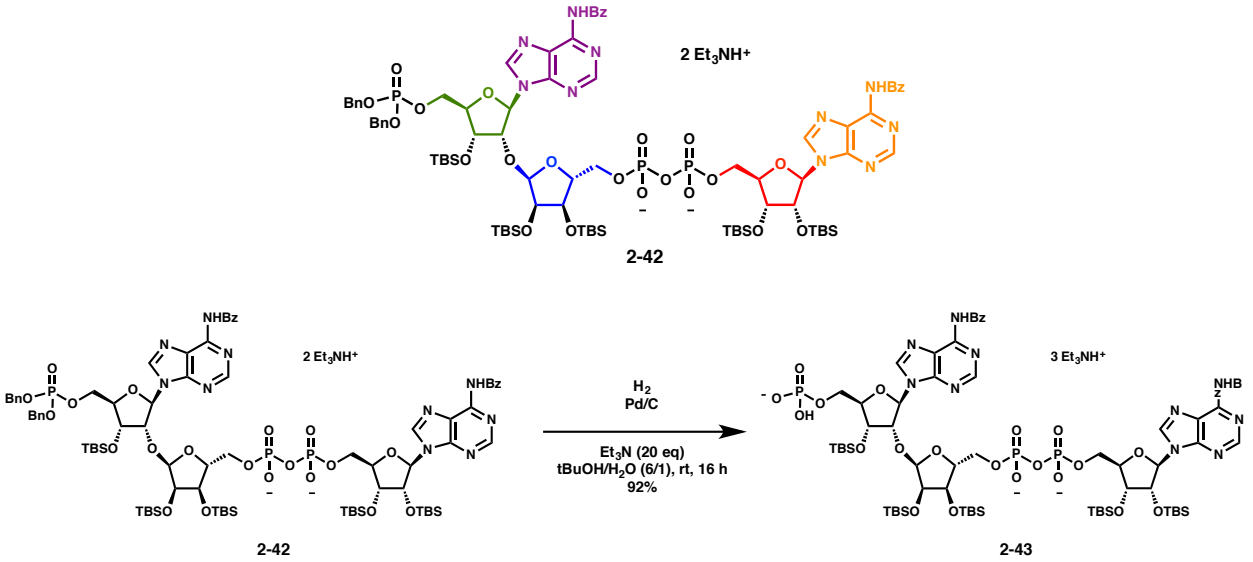
^{1}H NMR (500 MHz, CD_3OD) δ 8.96 (s, 1H), 8.71 (s, 1H), 8.66 (s, 1H), 8.45 (s, 1H), 8.12 – 8.00 (m, 4H), 7.71 – 7.60 (m, 2H), 7.59 – 7.48 (m, 4H), 7.37 – 7.19 (m, 10H), 6.27 (d, *J* = 2.7 Hz, 1H), 6.25 (d, *J* = 6.9 Hz, 1H), 5.20 (d, *J* = 4.3 Hz, 1H), 4.99 – 4.88 (m, 6H), 4.78 – 4.71 (m, 1H), 4.46 (dd, *J* = 4.3, 1.5 Hz, 1H), 4.40 – 4.29 (m, 3H), 4.28 – 4.13 (m, 6H), 4.12 – 4.03 (m, 2H), 3.16 (q, *J* = 7.3 Hz, 12H), 1.28 (t, *J* = 7.3 Hz, 18H), 0.98 (s, 9H), 0.92 (s, 9H), 0.91 (s, 9H), 0.81 (s, 9H), 0.73 (s, 9H), 0.19 (s, 3H), 0.17 (s, 3H), 0.16 (s, 3H), 0.12 (s, 3H), 0.10 (s, 3H), 0.09 (s, 3H), 0.04 (s, 3H), 0.02 (s, 3H), 0.01 (s, 3H), -0.36 (s, 3H).

^{13}C NMR (125 MHz, CD_3OD) δ 168.0, 168.0, 153.9, 153.4, 153.3, 152.9, 151.4, 151.3, 144.9, 144.5, 137.0, 137.0, 135.3, 135.1, 134.0, 134.0, 129.9, 129.8, 129.8, 129.6, 129.6, 129.2, 125.4, 124.8, 104.8, 90.1, 88.7, 87.4 (d, *J* = 8.9 Hz), 86.4 (d, *J* = 9.7 Hz), 83.5 (d, *J* = 8.0 Hz), 81.5, 77.6, 75.0, 74.7, 72.5, 71.8, 71.0, 71.0, 67.3 (d, *J* = 5.7 Hz), 66.7 (d, J_P = 5.8 Hz), 66.6 (d, J_P = 5.4 Hz), 47.4, 27.0, 26.8, 26.7, 26.7, 26.4, 19.3, 19.2, 19.2, 19.1, 18.8, 9.3, -3.6, -3.7, -3.9, -3.9, -4.0, -4.0, -4.1, -4.1, -4.5, -4.9.

^{31}P NMR (202 MHz, CD_3OD) δ -0.09, -10.01 (d, J = 19.9 Hz), -10.34 (d, J = 19.8 Hz).

HRMS (ESI) m/z calcd for $C_{83}H_{124}N_{10}O_{22}P_3Si_5$ ([M-H]$^-$) 1845.6956, found 1845.6959

	^{1}H	^{13}C	^{31}P	^{1}H-^{1}H COSY	^{1}H-^{13}C HMBC
I-Ado 1’	6.25 (d, J = 6.9 Hz)	88.69		4.89	74.73, 77.60, 87.43, 144.92, 153.89
I-Ado 2’	4.90 (dd, 6.9, 4.4 Hz)	77.60		4.47, 6.25	74.73, 88.69
I-Ado 3’	4.47 (d, J = 3.6 Hz)	74.73		4.25, 4.90	66.63, 77.60, 87.43, 88.69
I-Ado 4’	4.25	87.43 (d, J_P = 8.9 Hz)			
I-Ado 5’a	4.24	66.63	-10.34	4.35	
I-Ado 5’b	4.35	(d, J_P = 5.4 Hz)	(d, J = 19.8 Hz)	4.24	
I-Ade 2	8.71	153.39		-	124.81, 151.27, 153.89
I-Ade 4	-	153.89		-	-
I-Ade 5	-	124.81		-	-
I-Ade 6	-	151.27		-	-
I-Ade 8	8.95	144.92		-	124, 81, 153.89
a- Rib 1’’	5.20 (d, J = 4.2 Hz).	104.78		4.09	72.52, 75.03, 81.47, 86.40
a- Rib 2’’	4.09 (dd, J = 5.8, 4.2 Hz)	75.03		4.18, 5.20	104.78
a- Rib 3’’	4.18 (dd, J = 5.8, 3.1 Hz)	72.52		4.09	66.67, 75.03, 86.40
a- Rib 4’’	4.25	86.40 (d, J_P = 9.7 Hz).		4.07	75.03
a- Rib 5’’a	4.07	66.67	-10.01	4.25	72.52, 86.40
a- Rib 5’’b	4.07	(d, J_P = 5.8 Hz).	(d, J = 19.8 Hz)	4.25	72.52, 86.40
D-Ado 1’	6.27 (d, J = 2.6 Hz),	90.11		4.89	71.83, 81.47, 83.46, 144.53, 152.87
D-Ado 2’	4.90	81.47		4.74, 6.27	71.83, 83.46, 104.78
D-Ado 3’	4.74 (dd, J = 6.7, 4.5 Hz)	71.83		4.33, 4.90	67.25, 81.47, 83.46, 90.11
D-Ado 4’	4.33	83.46 (d, J_P = 8.0 Hz)			
D-Ado 5’a	4.20	67.25	-0.09	4.37	
D-Ado 5’b	4.37	(d, J_P = 5.7 Hz)		4.20	
D-Ade 2	8.66	153.32		-	125.37, 151.39, 152.87
D-Ade 4	-	152.87		-	-
D-Ade 5	-	125.37		-	-
D-Ade 6	-	151.39		-	-
D-Ade 8	8.44	144.53		-	125.37, 152.87



Compound 2-43: To a 20 mL vial was added compound **2-42** (51 mg, 0.025 mmol, 1 eq) and 10% palladium on activated charcoal (50 mg). Next, t-butanol (2.4 mL) and deionized water (0.4 mL) were added followed by Et_3N (0.070 mL, 0.498 mmol, 20 eq). Lastly, the vial was fitted with a balloon of hydrogen gas and purged for 1 min. The reaction was stirred under an atmosphere of hydrogen gas at room temperature for 16 h. The reaction was filtered through celite and the product was purified by reverse phase chromatography to yield compound **2-43** as a white foam (45 mg, 92%).

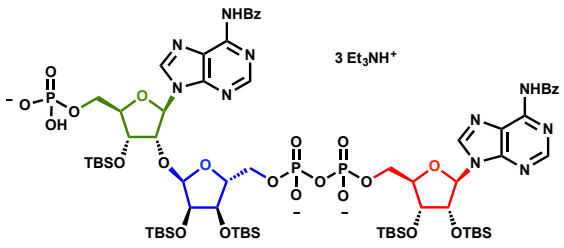
^{1}H NMR (600 MHz, CD_3OD) δ 9.00 (s, 1H), 8.84 (s, 1H), 8.71 (s, 1H), 8.70 (s, 1H), 8.12 – 8.03 (m, 4H), 7.67 – 7.62 (m, 2H), 7.60 – 7.52 (m, 4H), 6.34 (d, *J* = 3.6 Hz, 1H), 6.25 (d, *J* = 6.7 Hz, 1H), 5.26 (d, *J* = 4.1 Hz, 1H), 4.88 (dd, *J* = 6.8, 4.4 Hz, 1H), 4.67 (appt, *J* = 4.1 Hz, 1H), 4.61 (appt, *J* = 5.0 Hz, 1H), 4.46 (dd, *J* = 4.4, 1.7 Hz, 1H), 4.36-4.30 (m, 2H), 4.29 – 4.22 (m, 4H), 4.17 (dd, *J* = 5.7, 3.1 Hz, 1H), 4.11 – 4.02 (m, 4H), 3.18 (q, *J* = 7.3 Hz, 18H), 1.29 (t, *J* = 7.3 Hz, 27H), 0.98 (s, 9H), 0.93 (s, 9H), 0.91 (s, 9H), 0.82 (s, 9H), 0.74 (s, 9H), 0.19 (s, 3H), 0.17 (s, 3H), 0.16 (s, 3H), 0.15 (s, 3H), 0.08 (s, 3H), 0.08 (s, 3H), 0.03 (s, 3H), 0.03 (s, 3H), 0.02 (s, 3H), -0.33 (s, 3H).

^{13}C NMR (125 MHz, CD_3OD) δ 168.1, 168.1, 153.8, 153.8, 153.2, 153.2, 153.1, 153.1, 151.1, 151.0, 144.9, 144.3, 135.1, 133.8, 133.8, 129.7, 129.7, 129.4, 129.4, 125.0, 124.7, 104.9, 89.0, 88.7, 87.2 (d, *J* = 9.2 Hz), 85.9 (d, *J* = 9.8 Hz) 85.2 (d, *J* = 9.3 Hz) 82.8, 77.6, 74.7, 74.5, 72.3, 72.1, 66.5 (d, *J* = 5.5 Hz), 66.4 (d, *J* = 5.0 Hz), 64.6 (d, *J* = 4.6 Hz), 47.4, 26.9, 26.7, 26.7, 26.5, 26.3, 19.1, 19.1, 19.1, 18.9, 18.7, 9.2, -3.7, -3.8, -4.1, -4.1, -4.2, -4.2, -4.3, -4.3 -4.4, -5.0.

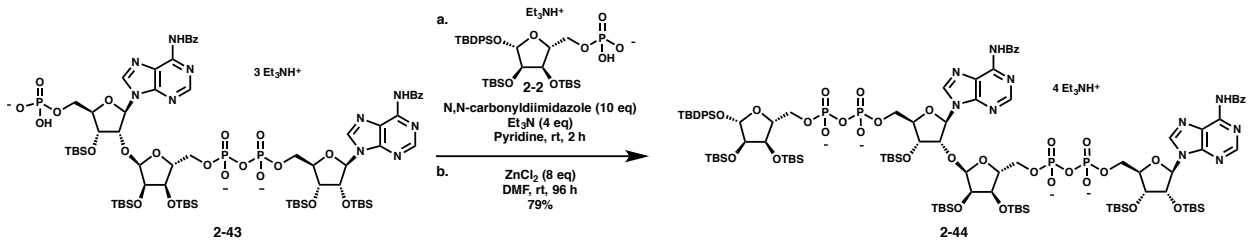
^{31}P NMR (202 MHz, CD_3OD) δ 1.86, -9.66 (d, *J* = 19.3 Hz), -10.06 (d, *J* = 19.1 Hz).

HRMS (ESI) m/z calcd for $C_{69}H_{112}N_{10}O_{22}P_3Si_5$ ($[M-H]^-$) 1665.6017, found 1665.6018

	^{1}H	^{13}C	^{31}P	^{1}H-^{1}H COSY
I-Ado 1’	6.25 (d, J = 6.7 Hz)	88.7		4.88
I-Ado 2’	4.88 (dd, J = 6.8, 4.4 Hz)	77.6		6.25, 4.46
I-Ado 3’	4.46 (dd, J =4.4, 1.7 Hz)	74.5		4.88, 4.25
I-Ado 4’	4.25	87.2 (d, J_P = 9.2 Hz)		4.46, 4.43, 4.25
I-Ado 5’a	4.25	66.4 (d, J_P = 5.0 Hz)	-10.06 (d, J = 19.1 Hz)	4.33, 4.25
I-Ado 5’b	4.33			4.25, 4.25
a- Rib 1’’	5.26 (d, J = 4.1 Hz)	104.9		4.04
a- Rib 2’’	4.04	74.7		5.26, 4.17
a- Rib 3’’	4.17	72.3		4.23,4.04
a- Rib 4’’	4.23	85.9 (d, J_P = 9.8 Hz)		4.23, 4.06, 4.06
a- Rib 5’’a	4.06	66.5 (d, J_P = 5.5 Hz)	-9.66 (d, J = 19.3 Hz)	4.23, 4.06
a- Rib 5’’b	4.06			4.23, 4.06
D-Ado 1’	6.34 (d, J = 3.6 Hz)	89.0		4.67
D-Ado 2’	4.67 (appt, J = 4.1 Hz)	82.8		6.34, 4.61
D-Ado 3’	4.61 (appt, J = 5.0 Hz)	72.1		4.67, 4.32
D-Ado 4’	4.32	85.2 (d, J_P = 9.3 Hz)		4.61, 4.26, 4.05
D-Ado 5’a	4.26	64.6 (d, J_P = 4.6 Hz)	1.86	4.32, 4.05
D-Ado 5’b	4.05			4.32, 4.26



2-43



Compound 2-44: To a 7 mL vial was added compound **2-2** (10 mg, 5.7 mmol, 1 eq). This material was azeotropically dried three times with anhydrous CH_3CN. To a separate 7 mL vial was added compound **3** (8.1 mg, 10.1 mmol, 2 eq) and carbonyldiimidazole (8.3 mg, 51.1 mmol, 10 eq). A solution of triethylamine (2.1 mg, 20.7 mmol, 4 eq) in pyridine (0.2 mL) was added and the reaction was stirred at room temperature for 2 h. The reaction was quenched with MeOH containing 10% Et_3N (v/v). After 30 min, the solution was concentrated to dryness and azeotropically dried with anhydrous CH_3CN (three times). Compound **2-43** was added in methanol to the crude imidazolide, the solution was concentrated to dryness and azeotropically dried with anhydrous CH_3CN (three times). The solid residue was dried under high vacuum over P_2O_5 for 16 h. A solution of freshly dried $ZnCl_2$ (5.5 mg, 40.4 mmol, 8 eq) in DMF (0.3 mL) was added and the reaction was stirred at room temperature for 96 h. The reaction was quenched by addition of a methanol solution containing the triethylammonium salt of EDTA (1 mL of methanol, 100 mg of EDTA (0.342 mmol, 67 eq, tetraacid), and 0.2 mL of triethylamine) and stirred for 30 min. The reaction was concentrated and purified by C18 chromatography to give compound **2-44** as the triethylammonium salt (11.0 mg, 79%).

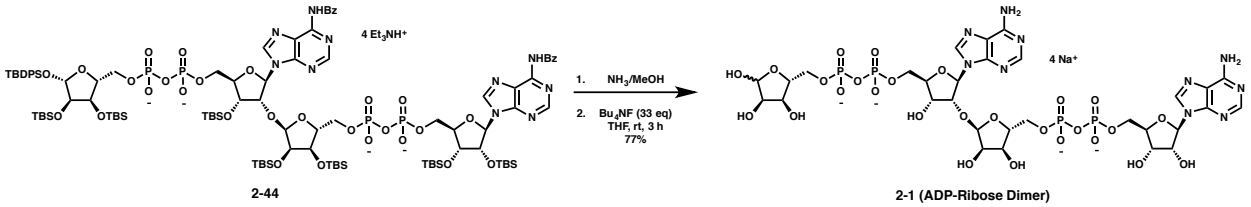
^{1}H NMR (600 MHz, CD_3OD) δ 9.01 (s, 1H), 8.97 (s, 1H), 8.71 (s, 1H), 8.70 (s, 1H), 8.10-8.05 (m, 4H), 7.76 – 7.69 (m, 2H), 7.68 – 7.61 (m, 4H), 7.56 – 7.52 (m, 4H), 7.47 – 7.35 (m, 6H), 6.33 (d, *J* = 3.7 Hz, 1H), 6.25 (d, *J* = 6.8 Hz, 1H), 5.23 (bs, 1H), 5.01 (s, 1H), 4.87 (m, 1H) 4.64

(bs, 1H), 4.59 (appt, J = 4.9 Hz, 1H), 4.46 (dd, J = 4.4, 1.6 Hz, 1H), 4.43-4.38 (m, 1H), 4.34-4.27 (m, 4H), 4.26 – 4.03 (m, 10H), 3.89 – 3.86 (m, 1H), 3.17 (q, J = 7.3 Hz, 24H), 1.28 (t, J = 7.3 Hz, 36H), 1.01 (s, 9H), 0.97 (s, 9H), 0.91 (s, 9H), 0.90 (s, 9H), 0.90 (s, 9H) 0.79 (s, 9H), 0.76 (s, 9H), 0.74 (s, 9H), 0.18 (s, 3H), 0.18 (s, 3H), 0.16 (s, 3H), 0.15 (s, 3H), 0.14 (s, 3H), 0.13 (s, 3H), 0.13 (s, 3H), 0.11 (s, 3H), 0.07 (s, 3H), 0.07 (s, 3H), 0.01 (s, 3H), -0.08 (s, 3H), -0.29 (s, 3H), -0.34 (s, 3H).

^{13}C NMR (125 MHz, CD_3OD) δ 168.1, 168.0, 153.8, 153.2, 153.2, 153.1, 151.1, 151.0, 144.9, 144.4, 137.0, 136.8, 135.7, 135.2, 135.1, 134.6, 133.9, 133.8, 131.2, 131.0, 129.8, 129.4, 129.0, 129.0, 128.9, 128.9, 124.9, 124.6, 105.0, 102.9, 88.8, 88.6, 87.3, 87.2, 86.2, 85.2, 83.1, 82.2, 79.2, 77.7, 74.8, 74.6, 74.5, 72.3, 69.8, 66.6, 66.4, 65.3, 47.3, 27.4, 27.0, 26.8, 26.7, 26.7, 26.5, 26.3, 26.3, 19.9, 19.1, 19.1, 19.1, 18.9, 18.8, 18.8, 18.7, 9.2, -3.7, -3.8, -3.9, -4.0, -4.1, -4.2, -4.2, -4.2, -4.2, -4.2, -4.3, -4.3, -4.4, -5.0.

^{31}P NMR (243 MHz, CD_3OD) δ -10.84 (d, J = 18.5 Hz), -11.01 (d, J = 19.5 Hz), -11.39 (d, J = 19.3 Hz), -11.44 (*d*, J = 18.3 Hz).

HRMS (ESI) m/z calcd for $C_{102}H_{167}O_{29}Si_8P_4$ ($[M-H]^-$) 2343.9011, found 2343.9024.



2-44 2-1 (ADP-Ribose Dimer)

Compound 2-1: To a 4 mL reaction vial was added compound **2-44** (9.0 mg, 3.27 mmol, 1 eq) and 7M ammonia in methanol (1.5 mL). The reaction was at room temperature for 20 h at which point the solution was concentrated to dryness. The residue was dissolved in methanol containing 10% triethylamine (v/v) and concentrated to dryness (twice). The solution was then

azeotropically dried with CH_3CN (three times), and dried over P_2O_5 under high vacuum for 2 h. Next, THF (0.22 mL) was added followed by tetrabutylammonium fluoride (0.1 mL of a 1 M solution in THF, 0.11 mmol, 33 eq). The reaction was stirred for 3 h at room temperature. The reaction was concentrated to a total volume of ~0.1 mL, 0.1 mL of a 3M aqueous NaOAc solution was added, and the solution was stirred for 0.5 h. The solution was transferred in 0.1 mL aliquots to 1.7 mL centrifuge tubes and 0.95 mL of ethanol was added to each tube. The tubes were cooled to -78^{o}C for 0.5 h and were centrifuged at 4^{o}C for 20 min at 14,000 rcf. The supernatant was discarded and the pellet was resuspended in 1 mL of fresh ethanol. The tubes were again cooled at -78^{o}C and centrifuged as described above. The pellet was dissolved in 1 mL of water and purified by reverse phase chromatography (99% 8 mM Et_3NHOAc/1% CH_3CN to 90% 8 mM Et_3NHOAc/10% CH_3CN over 12 min). The fractions containing the product were concentrated to dryness followed by repeated lyophillization with H_2O and exchange for the sodium salt (Dowex 50W-X8 Na^+ form) to produce compound **2-1** (3.0 mg, 77%) as the sodium salt. Note: NMR Data for compound **2-1** was obtained before exchange for the Na^+ ion and contains ~10 eq of Et_3NHOAc. It was observed that removal of the buffer beyond this point caused dramatic suppression of the NMR signal, which we believe is due to a compound aggregation phenomenon. Addition of a solution containing ~10 eq Et_3NHOAc in D_2O to buffer-free samples of **2-1** restored its NMR signal.

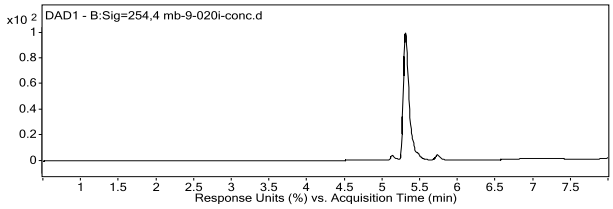
^{1}H NMR (600 MHz, D_2O) δ 8.47 (bs, 2H), 8.24 (bs, 2H), 6.22 (d, *J* = 3.1 Hz, 1H), 6.03 (d, *J* = 5.6 Hz, 1H), 5.35 (d, *J* = 4.1 Hz, 1H), 5.33 (d, *J* = 4.3 Hz, 1H), 5.23 (d, J = 2.2 Hz, 1H), 4.71 (appt, *J* = 5.7 Hz, 1H), 4.65-4.61 (m, 1H), 4.60 – 4.56 (m, 1H), 4.52 – 4.49 (m, 1H), 4.43 – 4.38 (m, 1H), 4.38 – 4.33 (m, 3H), 4.33 – 4.30 (m, 1H), 4.29 – 4.17 (m, 7H), 4.13 – 4.01 (m, 9H).

^{13}C NMR (125 MHz, D_2O) δ See table; assignments and shifts from HSQC/HMBC.

^{31}P NMR (243 MHz, D_2O) δ -9.65 - -11.98 (m, 4P)

HRMS (ESI) m/z calcd for $C_{30}H_{43}N_{10}O_{27}P_4$ ([M-H]$^-$) 1099.1255 found 1099.1221.

LCMS: tR = 5.3 min, 98% 5 mM pentylamine:HOAc (pH=6.5) with 2% CH_3CN to 75% 6 mM pentylamine:HOAc (pH=6.5) with 25% CH_3CN over 5 min, then hold 75% 6 mM pentylamine:HOAc (pH=6.5) with 25% CH_3CN for 3 min



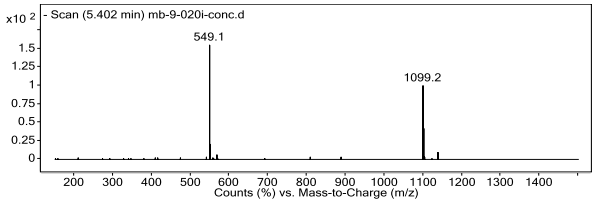
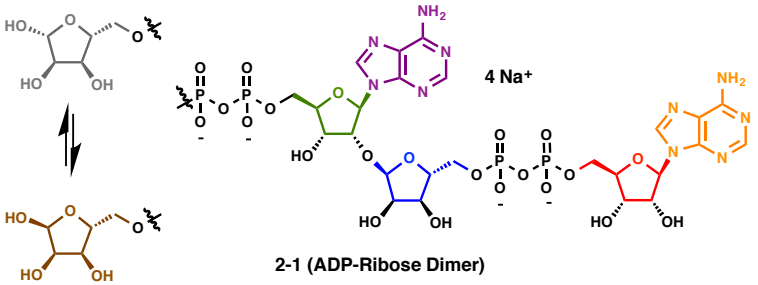
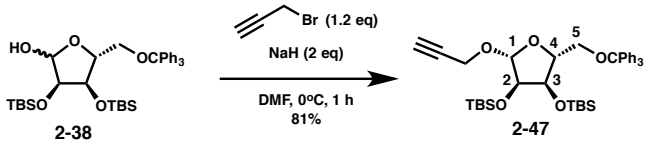


Table of NMR Data:

	^{1}H	^{13}C	^{1}H-^{1}H COSY	^{1}H-^{13}C HMBC
I-Ado 1’	6.03 (d, *J* = 5.9 Hz)	87.23	4.71	74.63, 140.45, 148.3
I-Ado 2’	4.71 (at, 5.4 Hz)	74.63	6.03, 4.51	84.02, 87.23
I-Ado 3’	4.51 (dd, *J* = 4.9, 3.3 Hz)	70.63	4.71, 4.37	65.01, 87.25
I-Ado 4’	4.36	84.02	4.51	65.01
I-Ado 5’a	4.22	65.01		
I-Ado 5’b	4.22			
I-Ade 2	8.25	150.45	-	118.24, 148.33, 153.81
I-Ade 4	-	148.33	-	-
I-Ade 5	-	118.24	-	-
I-Ade 6	-	153.81	-	-
I-Ade 8	8.47	140.45	-	118.49, 148.36
D-Rib 1’’	5.33 (d, *J* = 4.3 Hz)	101.42		69.88, 71.67, 78.92, 84.31,
D-Rib 2’’	4.27	71.48		84.52
D-Rib 3’’	4.23	69.94		84.31, 65.67
D-Rib 4’’	4.36	84.31		
D-Rib 5’’a	4.07	65.8		
D-Rib 5’’b	4.07			
D-Ado 1’	6.22 (d, *J* = 3.1 Hz)	87.04	4.63	140.45
D-Ado 2’	4.63 (m)	78.94	6.22, 4.58	101.34, 83.21
D-Ado 3’	4.58 (m)	68.94	4.63, 4.40	64.52, 86.94
D-Ado 4’	4.40	83.15	4.58	
D-Ado 5’a	4.35	64.49		68.94
D-Ado 5’b	4.25*			
D-Ade 2	8.25	150.45	-	118.24, 148.33, 153.81
D-Ade 4	-	148.33	-	-
D-Ade 5	-	118.24	-	-

D-Ade 6	-	153.81	-	-
D-Ade 8	8.47	140.45	-	118.49, 148.36
a Rib 1''	5.35 (d, *J* = 4.1 Hz)	96.40	4.13	70.09, 81.87
a Rib 2''	4.14	70.81		
a Rib 3''	4.19	70.11		96.36, 65.67
a Rib 4''	4.24	81.92		
a Rib 5''a	4.07	65.8		
a Rib 5''b	4.04			
b Rib 1''	5.23 (d, *J* = 2.2 Hz)	101.23	4.03	70.72, 81.26
b Rib 2''	4.03	75.22	4.32, 5.23	81.16
b Rib 3''	4.32	70.55	4.03	66.38, 101.28
b Rib 4''	4.10	81.19		
b Rib 5''a	4.13	66.41		
b Rib 5''b	4.07			





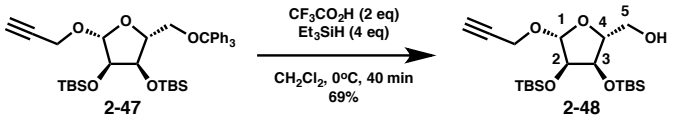
Compound 2-47: To a 0°C solution of compound **2-38** (90 mg, 0.145 mmol, 1 eq) in DMF (1 mL) was added sodium hydride (60% dispersion in mineral oil, 12.8 mg, 0.322 mmol, 2.2 eq) followed by propargyl bromide (0.027 mL, 0.242 mmol, 1.7 eq). The solution was stirred at 0°C for 1 h at which time it was poured into saturated aqueous NH_4Cl and extracted three times with diethyl ether. The combined ether extracts were washed three times with water, dried through Na_2SO_4, concentrated and purified by silica column chromatography to afford compound **2-47** (80 mg, 81%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.55 – 7.49 (m, 6H, -CPh_3), 7.33 – 7.27 (m, 6H, -CPh_3), 7.26 – 7.20 (m, 3H), 5.12 (s, 1H, H-1), 4.33 (dd, *J* = 2.4, 1.6 Hz, 2H, -CH_2CCH), 4.24 – 4.17 (m, 2H, H-3, H-4), 3.98 (d, *J* = 2.8 Hz, 1H, H-2), 3.37 (dd, *J* = 10.2, 1.7 Hz, 1H, H-5a), 3.09 – 3.00 (m,

1H, H-5b), 2.44 (t, *J* = 2.4 Hz, 1H, -CH_2CCH), 0.93 (s, 9H, -tBu), 0.72 (s, 9H, -tBu), 0.13 (s, 3H, -CH_3), 0.09 (s, 3H-CH_3), -0.05 (s, 3H-CH_3), -0.25 (s, 3H-CH_3).

^{13}C NMR (126 MHz, $CDCl_3$) δ 144.1, 128.9, 127.9, 127.0, 105.3, 86.4, 81.5, 79.4, 76.3, 74.6, 71.9, 64.0, 54.5, 25.9, 25.9, 18.2, 18.0, -4.1, -4.3, -4.6, -5.1.

HRMS (ESI) m/z calcd for $C_{39}H_{54}O_5Si_2Na$ ($[M+Na]^+$) 681.3408, found 681.3408.

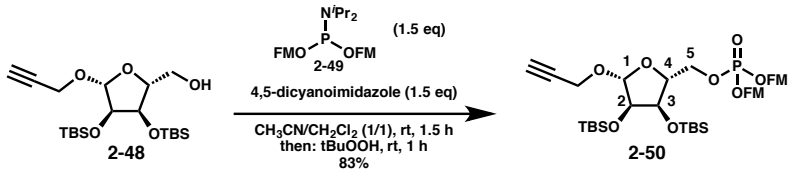


Compound 2-48: To a 0°C solution of compound **2-47** (80 mg, 0.121 mmol, 1 eq) in CH_2Cl_2 (1 mL) was added triethylsilane (0.077 mL, 0.486 mmol, 1 eq) followed by a solution of trifluoroacetic acid (28 mg, 0.243 mmol, 2 eq) in CH_2Cl_2 (0.1 mL) and the reaction was stirred at 0°C for 40 min. The reaction was quenched with saturated aqueous $NaHCO_3$ (1 mL), extracted three times with chloroform, dried through Na_2SO_4, concentrated and purified by silica column chromatography to afford compound **2-48** (35 mg, 69%) as a white foam.

^{1}H NMR (500 MHz, $CDCl_3$) δ 4.95 (s, 1H, H-1), 4.33 – 4.18 (m, 3H, -CH_2CCH, H-3), 4.10 – 4.03 (m, 1H, H-4), 3.95 (d, *J* = 4.0 Hz, 1H, H-2), 3.85 (dd, *J* = 12.2, 2.5 Hz, 1H, H-5a), 3.57 (dd, *J* = 12.2, 3.5 Hz, 1H, H-5b), 2.46 (t, *J* = 2.4 Hz, 1H, -CH_2CCH), 1.88 (bs, 1H, -OH), 0.91 (s, 9H, -tBu), 0.89 (s, 9H, -tBu), 0.10 (s, 6H, -CH_3), 0.09 (s, 3H, -CH_3), 0.08 (s, 3H, -CH_3).

^{13}C NMR (126 MHz, $CDCl_3$) δ 106.6, 82.8, 79.7, 76.7, 74.7, 70.6, 61.5, 55.4, 26.0, 25.9, 18.2, 18.2, -4.1, -4.4, -4.5, -4.9.

HRMS (ESI) m/z calcd for $C_{20}H_{41}O_5Si_2$ ($[M+H]^+$) 417.2493, found 417.2489.



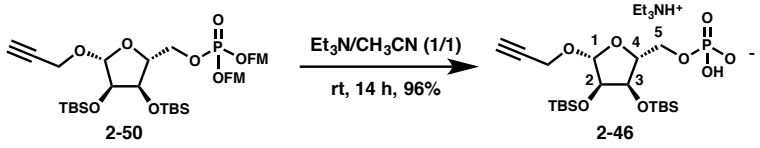
Compound 2-50: To a room temperature solution of compound **2-48** (10 mg, 0.024 mmol, 1 eq) in CH_2Cl_2 (0.3 mL) was added phosphoramidite **2-49** (18.8 mg, 0.036 mmol, 1.5 eq) in CH_2Cl_2 (0.1 mL) followed by 4,5-dicyanoimidazole (4.2 mg, 0.036 mmol, 1.5 eq) in CH_3CN (0.4 mL). The reaction was stirred at room temperature for 1.5 h, at which point tBuOOH (0.021 mL of 5.5 M solution, 0.12 mmol, 5 eq) was added and the reaction was stirred for an additional 1 h. The reaction was poured into saturated aqueous $NaHCO_3$, extracted three times with chloroform, dried through Na_2SO_4, concentrated and purified by silica column chromatography to give compound **2-50** (17 mg, 83%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.79 – 7.67 (m, 4H, -FM), 7.60 – 7.48 (m, 4H, -FM), 7.43 – 7.31 (m, 4H, -FM), 7.30 – 7.20 (m, 4H, -FM), 4.95 (s, 1H, H-1), 4.32-4.24 (m, 4H, -C*H*$_2$CHAryl), 4.22 – 4.03 (m, 7H, -CH$_2$C*H*Aryl + H-4 + H-5a + H-3, -C*H*$_2$CCH), 3.96 (appdt, *J* = 10.6, 5.1 Hz, 1H, H-5b), 3.92 (d, *J* = 3.8 Hz, 1H, H-2), 2.28 (t, *J* = 2.4 Hz, 1H, -CH$_2$CC*H*), 0.89 (s, 9H, -tBu), 0.86 (s, 9H, -tBu), 0.09 (s, 3H, -CH$_3$), 0.08 (s, 3H, -CH$_3$), 0.03 (s, 3H, -CH$_3$), 0.00 (s, 3H, -CH$_3$).

^{13}C NMR (126 MHz, $CDCl_3$) δ 143.3, 141.5, 128.0, 127.2, 125.3, 120.1, 104.9, 80.4 (d, *J* = 7.8 Hz), 79.2, 76.3, 74.7, 71.4, 69.5 (appt, *J* = 6.3 Hz), 67.2 (d, *J* = 5.2 Hz), 54.1, 48.1 (d, *J* = 8.1 Hz), 25.9, 25.9, 18.2, 18.1, -4.1, -4.4, -4.5, -5.0.

^{31}P NMR (202 MHz, CD_3OD) δ -0.80

HRMS (ESI) m/z calcd for $C_{48}H_{62}O_8PSi_2$ ($[M+H]^+$) 853.3721, found 853.3718.



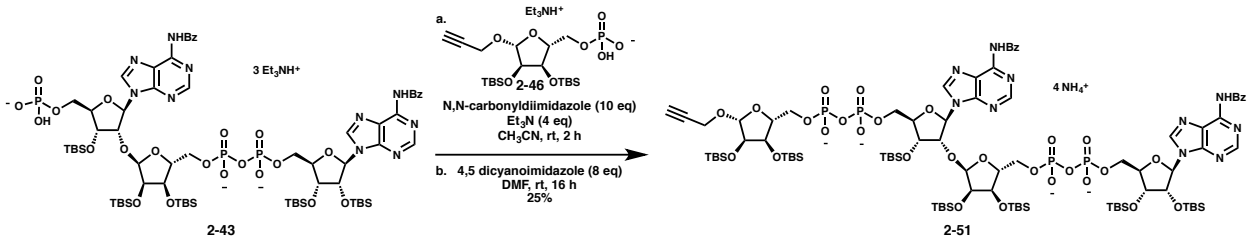
Compound 2-46: To a room temperature solution of compound **2-50** (17 mg, 0.020 mmol, 1 eq) in CH_3CN (0.4 mL) was added Et_3N (0.2 mL). The solution was stirred for 14 h at which point it was evaporated and purified by C-18 chromatography to give compound **2-46** (11.5 mg, 96%).

^{1}H NMR (500 MHz, CD_3OD) δ 5.05 (d, *J* = 1.3 Hz, 1H, H-1), 4.34 (dd, *J* = 16.1, 2.4 Hz, 1H, -C*H*$_2$CCH), 4.28 – 4.20 (m, 2H, -C*H*$_2$CCH + H-3), 4.10 – 4.02 (m, 2H, H-4 + H-5a), 3.98 (dd, *J* = 4.1, 1.3 Hz, 1H, H-2), 3.81 (ddd, *J* = 11.3, 6.7, 4.2 Hz, 1H, H-5b), 3.19 (q, *J* = 7.3 Hz, 6H, (CH_3CH_2)$_3$N), 2.81 (appt, *J* = 2.4 Hz, 1H, -CH_2CCH), 1.31 (t, *J* = 7.3 Hz, 9H(CH_3CH_2)$_3$N), 0.93 (s, 9H, -tBu), 0.92 (s, 9H, -tBu), 0.14 (s, 3H, -CH_3), 0.13 (s, 3H, -CH_3), 0.12 (s, 3H, -CH_3), 0.11 (s, 3H, -CH_3).

^{13}C NMR (126 MHz, CD_3OD) δ 105.5, 83.5 (d, *J* = 9.0 Hz), 80.2, 77.8, 75.7, 73.5, 66.5 (d, *J* = 5.6 Hz) 54.7, 47.8, 26.5, 26.4, 19.0, 18.9, 9.2, -4.1, -4.3, -4.3, -4.6.

^{31}P NMR (202 MHz, CD_3OD) δ 2.11

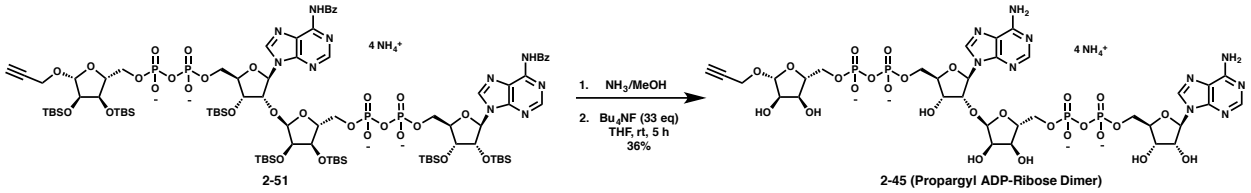
HRMS (ESI) m/z calcd for $C_{20}H_{40}O_8PSi_2$ ([M-H]$^-$) 495.2005, found 495.1999.



Compound 2-51: To a 4 mL reaction vial was added compound **2-46** (4.3 mg, 7.2 mmol, 2 eq) and the compound was azeotroped three times with CH_3CN and dried under high vacuum. Compound **2-46** was dissolved in CH_3CN (0.2 mL) and Et_3N (1.5 mg, 14.4 mmol, 4 eq). N,N-carbonyldiimidazole (5.8 mg, 36 mmol, 10 eq) was added and the reaction stirred for 2 h at room temperature. Excess carbonydiimidazole was quenched using 10% Et_3N in methanol (0.2 mL) and the crude reaction was concentrated to dryness. Compound **2-43** (7.0 mg, 3.6 mmol, 1 eq)

was added and the solution was azeotroped three times with CH_3CN and dried under high vacuum. The coupling reaction was initiated by addition of of 4,5-dicyanoimidazole (3.4 mg, 29 mmol, 8 eq) in DMF (0.2 mL) and was stirred for 16 h at room temperature. The crude reaction was applied to a column of silica and eluted using a gradient of 90:0:10 iPrOH:H_2O:NH_4OH to 70:20:10 iPrOH:H_2O:NH_4OH to give compound **2-451** as the ammonium salt (2.0 mg, 25%) and an inseparable by-product resulting from dimerization of compound **2-46**.

HRMS (ESI) m/z calcd for $C_{89}H_{152}N_{10}O_{29}P_4Si_7$ ([M-H]$^-$) 2144.8012, found 2144.7959.



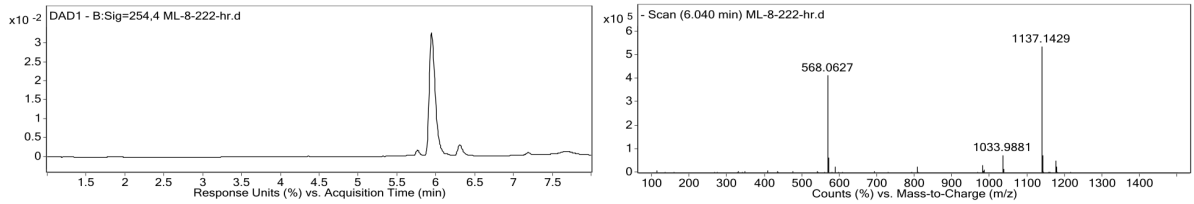
Compound 2-45: To a 4 mL reaction vial was added compound **2-51** (2.0 mg, 0.903 mmol, 1 eq) and NH_3/MeOH (0.2 mL of a 7M solution). The solution was stirred for 14 h, the solvent was evaporated and 0.5 mL of 40% Et_3N in MeOH was added and subsequently evaporated. A solution of Bu_4NF (0.2 mL of 1M solution in THF) was added and the reaction was stirred for 5 h at room temperature. To the solution was added 0.2 mL of a 3M aqueous NaOAc solution and the solution stirred for 20 min. Ethanol (0.5 mL) was added, the solution was aliquoted into centrifuge tubes and put into a -80^{o}C freezer. The tubes were centrifuged for 5 min at 14,000 rcf, the supernatant was discarded, ethanol was again added and the tubes were again centrifuged. The pellet was dissolved in water and purified by preparative HPLC (99% 8 mM Et_3NHOAc/1% CH_3CN to 90% 8 mM Et_3NHOAc/10% CH_3CN over 12 min). The fractions containing the product were concentrated to dryness followed by repeated lyophillization with H_2O and exchange for the ammonium salt (Dowex 50W-X8 NH_4^+ form) to produce compound **2-45** (0.4 mg, 36%) as the ammonium salt.

^{1}H NMR (600 MHz, D_2O) δ 8.41 (s, 1H), 8.39 (s, 1H), 8.17 (s, 1H), 8.16 (s, 1H), 6.19 (d, *J* = 3.3 Hz, 1H), 6.02 (d, *J* = 6.0 Hz, 1H), 5.31 (d, *J* = 4.2 Hz, 1H), 5.10 (s, 1H), 4.68 (appt, *J* = 5.6 Hz, 1H), 4.65 (q, *J* = 2.5 Hz, 1H), 4.61 (appt, *J* = 5.7 Hz, 1H), 4.52 – 4.48 (m, 1H), 4.42 – 4.17 (m, 13H), 4.17-4.06 (m, 3H), 4.03 (d, *J* = 4.6 Hz, 1H), 4.00-3.94 (m, 1H), 2.82 (s, 1H).

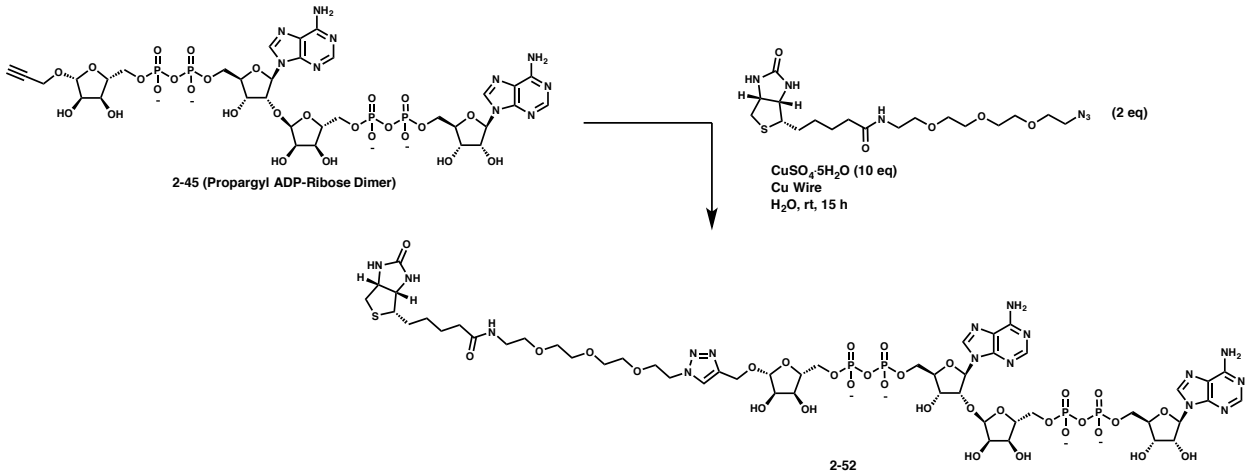
^{31}P NMR (243 MHz, D_2O) δ -10.55 – -11.90 (m, 4P).

HRMS (ESI) m/z calcd for $C_{33}H_{46}N_{10}O_{27}P_4$ ([M-H]$^-$) 1137.1412, found 1137.1423.

LCMS: tR = 6.0 min, 98% 5 mM pentylamine:HOAc (pH=6.5) with 2% CH_3CN to 75% 6 mM pentylamine:HOAc (pH=6.5) with 25% CH_3CN over 5 min, then hold 75% 6 mM pentylamine:HOAc (pH=6.5) with 25% CH_3CN for 3 min



Note: m/z = 1033.9 corresponds to a component of the reference mixture

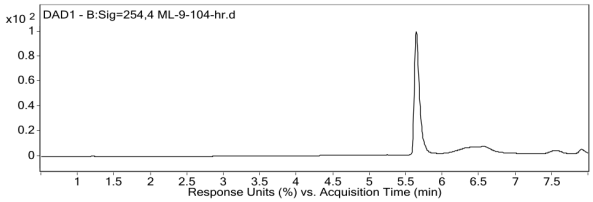


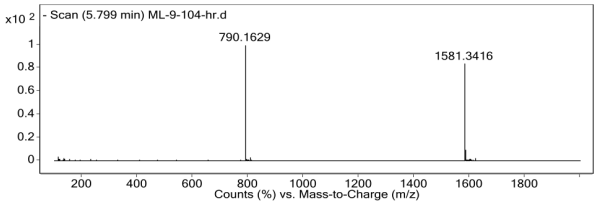
Compound 2-52: To a well of a 384 well plate was added a solution of compound **2-45** (10 mL of a 2 mM aqueous solution, 0.002 mmol, 1 eq), a solution of Peg_3-biotin-N_3 (10 mL of a 4 mM aqueous solution, 0.004 mmol, 2 eq), a solution of $CuSO_4{\cdot}5H_2O$ (10 mL of a 20 mM aqueous

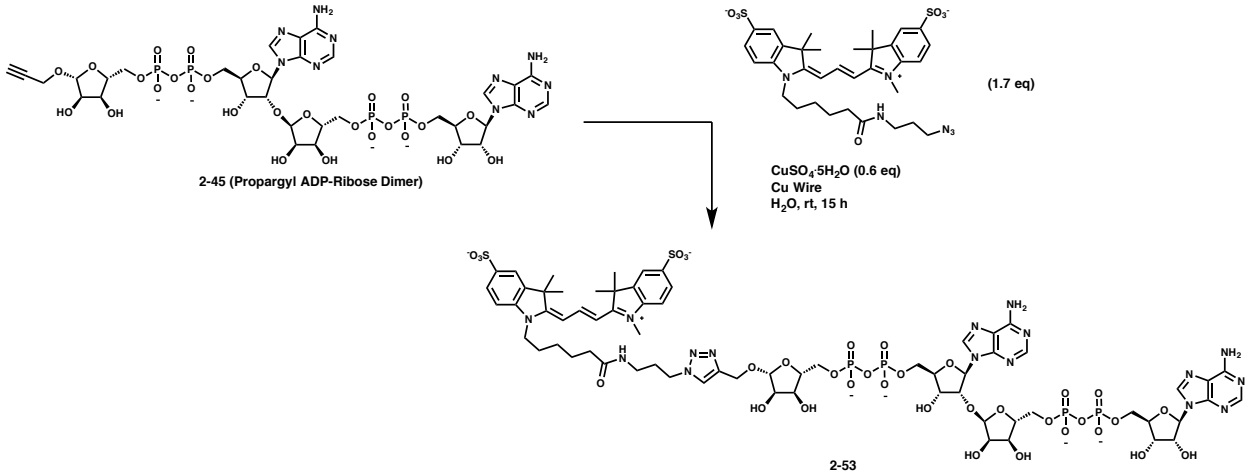
solution) and copper wire (3.2 mg). The solution was diluted to a final volume of 50 mL and incubated at room temperature overnight. The solution was purified by reverse phase HPLC (99% 8 mM Et_3NHOAc/1% CH_3CN to 70% 8 mM Et_3NHOAc/30% CH_3CN over 12 min) to give compound **2-52**. Compound **2-52** was >95% pure by LC/MS analysis.

HRMS (ESI) m/z calcd for $C_{51}H_{78}N_{16}O_{32}P_4S$ ($[M-H]^-$) 1581.3566, found 1581.3615.

LCMS: tR = 5.80 min, 98% 5 mM pentylamine:HOAc (pH=6.5) with 2% CH_3CN to 65% 6 mM pentylamine:HOAc (pH=6.5) with 35% CH_3CN over 5 min, then hold 65% 6 mM pentylamine:HOAc (pH=6.5) with 35% CH_3CN for 3 min







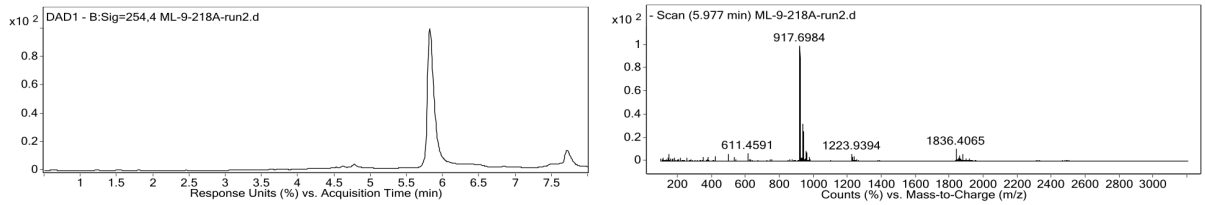
Compound 2-53: To a 500 mL eppendorf tube was added copper wire (2 mg). A solution of compound **2-45** (50 mL of a 3.1 mM solution, 0.16 mmol, 1 eq), a solution of sulfoCy3-N_3 (20 mL of a 10 mg/mL solution, 0.27 mmol, 1.7 eq) and a solution of $CuSO_4 \cdot 5H_2O$ (50 mL of a 0.5 mg/mL solution, 0.1 mmol, 0.6 eq). The reaction was incubated at room temperature for 15 h and

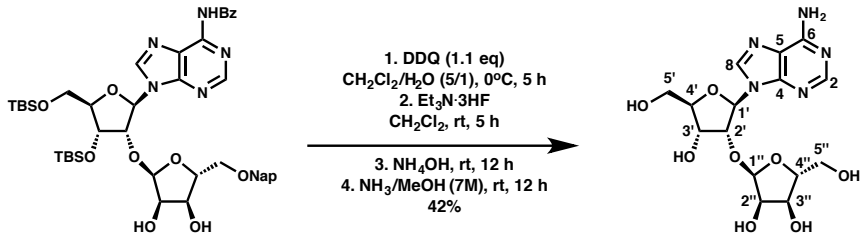
purified by HPLC (99% 8 mM Et_3NHOAc/1% CH_3CN to 70% 8 mM Et_3NHOAc/30% CH_3CN over 12 min) to give compound **2-53**. Compound **2-53** was >95% pure by LC/MS analysis.

HRMS (ESI) m/z calcd for $C_{66}H_{89}N_{16}O_{34}P_4S_2$ ($[M-H]^-$) 1836.4046, found 1836.4058.

m/z calcd for $C_{66}H_{89}N_{16}O_{34}P_4S_2$ ($[M-2H]^2$) 917.6987, found 917.6983.

LCMS: tR = 5.97 min, 98% 5 mM pentylamine:HOAc (pH=6.5) with 2% CH_3CN to 65% 6 mM pentylamine:HOAc (pH=6.5) with 35% CH_3CN over 5 min, then hold 65% 6 mM pentylamine:HOAc (pH=6.5) with 35% CH_3CN for 3 min





2-23 **2-22**

Compound 2-22: To a 4 mL reaction vial was added compound **2-23** (50 mg, 0.057 mmol, 1 eq), CH_2Cl_2 (1.3 mL), and H_2O (0.26 mL). The solution was cooled to 0 ^oC, and 2,3-Dichloro-5,6-dicyano-1,4-benzoquinone (14.3 mg, 0.063 mmol, 1.1 eq) was added. The reaction was stirred at 0 ^oC for 3 h, at which point an additional 1.1 eq of 2,3-Dichloro-5,6-dicyano-1,4-benzoquinone was added and the reaction stirred for 1 h. The reaction was quenched with a saturated aqueous $Na_2S_2O_3$ solution, extracted three times with EtOAc, filtered through Na_2SO_4 and concentrated. The crude reaction mixture was dissolved in CH_2Cl_2 (0.2 mL), $Et_3N{\cdot}3HF$ (46 mL, 0.285 mmol, 5 eq) was added, and the reaction stirred at room temperature for 5 h. Next, NH_4OH (0.16 mL of a 25% solution) was added and the reaction stirred for 12 h. The crude

reaction was poured into CH_2Cl_2, extracted three times with water and concentrated. Silica column chromatography was performed revealing two compounds (Rf = 0.38 and 0.15 in 80:20 CH_2Cl_2:MeOH). The mixture was treated with NH_3/MeOH (0.3 mL of a 7 M solution) for 12 h. Next, Dowex 50W-X8 (Na^+ form) was added and stirred for 15 min. The mixture was filtered through a cotton plug, concentrated and purified by silica column chromatography (R_f 0.15 in 80:20 CH_2Cl_2:MeOH) to furnish compound **2-22** as a yellow oil (9.5 mg, 42%).

^{1}H NMR (500 MHz, D_2O) δ 8.33 (s, 1H, H-2), 8.20 (s, 1H, H-8), 6.19 (d, *J* = 6.2 Hz, 1H, H-1'), 5.07 (d, *J* = 4.1 Hz, 1H, H-1''), 4.88 (dd, *J* = 6.3, 5.2 Hz, 1H, H-2'), 4.54 (dd, *J* = 5.2, 3.0 Hz, 1H, H-3'), 4.32 (appq, *J* = 3.1 Hz, 1H, H-4'), 4.20 (appdt, *J* = 4.9, 3.4 Hz, 1H, H-4''), 4.10 – 4.02 (m, 2H, H-2'' + H-3''), 3.92 (dd, *J* = 12.9, 2.7 Hz, 1H, H-5'a), 3.84 (dd, *J* = 12.9, 3.5 Hz, 1H, H-5'b), 3.68 (dd, *J* = 12.5, 3.6 Hz, 1H, H-5''a), 3.61 (dd, *J* = 12.5, 4.8 Hz, 1H, H-5''b).

^{13}C NMR (125 MHz, D_2O) δ 156.3, 153.2, 149.0, 141.4, 119.8, 102.6, 87.7, 86.9, 86.1, 79.8, 72.2, 71.6, 70.4, 62.2, 62.1.

HRMS (ESI) m/z calcd for $C_{15}H_{22}N_5O_8$ ($[M+H]^+$) 400.1477, found 400.1468.

2.5 References

1. Krietsch, J. *et al.* Reprogramming cellular events by poly(ADP-ribose)-binding proteins. *Mol Aspects Med* **34,** 1066–1087 (2013).
2. Gagne, J. P. *et al.* Proteome-wide identification of poly(ADP-ribose) binding proteins and poly(ADP-ribose)-associated protein complexes. *Nucleic Acids Res.* **36,** 6959–6976 (2008).
3. Pleschke, J. M., Kleczkowska, H. E., Strohm, M. & Althaus, F. R. Poly(ADP-ribose) binds to specific domains in DNA damage checkpoint proteins. *J. Biol. Chem.* **275,** 40974–40980 (2000).
4. Ahel, I. *et al.* Poly(ADP-ribose)-binding zinc finger motifs in DNA repair/checkpoint proteins. *Nature* **451,** 81–85 (2008).
5. Fahrer, J., Kranaster, R., Altmeyer, M., Marx, A. & Burkle, A. Quantitative analysis of the binding affinity of poly(ADP-ribose) to specific binding proteins as a function of chain length. *Nucleic Acids Res.* **35,** e143–e143 (2007).
6. Wang, Z. *et al.* Recognition of the iso-ADP-ribose moiety in poly(ADP-ribose) by WWE

domains suggests a general mechanism for poly(ADP-ribosyl)ation-dependent ubiquitination. *Genes Dev.* **26,** 235–240 (2012).
7. Oberoi, J. *et al.* Structural Basis of Poly(ADP-ribose) Recognition by the Multizinc Binding Domain of Checkpoint with Forkhead-associated and RING Domains (CHFR). *J. Biol. Chem.* **285,** 39348–39358 (2010).
8. Barkauskaite, E. *et al.* Visualization of poly(ADP-ribose) bound to PARG reveals inherent balance between exo- and endo-glycohydrolase activities. *Nat. Commun.* **4,** (2013).
9. Popp, O. *et al.* Site-Specific Noncovalent Interaction of the Biopolymer Poly(ADP-ribose) with the Werner Syndrome Protein Regulates Protein Functions. *ACS Chem. Biol.* **8,** 179–188 (2013).
10. Imai, J. & Torrence, P. F. Bis(2,2,2-trichloroethyl) phosphorochloridite as a reagent for the phosphorylation of oligonucleotides: preparation of 5'-phosphorylated 2“,5”-oligoadenylates. *J. Org. Chem.* **46,** 4015–4021 (1981).
11. Markiewicz, W. T. & Wiewiorowski, M. A new type of silyl protecting groups in nucleoside chemistry. *Nucleic Acids Res.* **1,** s185–s190 (1978).
12. Hakimelahi, G. H., Proba, Z. A. & Ogilvie, K. K. High yield selective 3′-silylation of ribonucleosides. *Tetrahedron Lett.* **22,** 5243–5246 (1981).
13. Cole, A. G. & Gani, D. 'Active“ conformation of the inositol monophosphatase substrate, adenosine 2”-phosphate: role of the ribofuranosyl O-atoms in chelating a second Mg2+ ion. *J. Chem. Soc., Perkin Trans. 1* 2685–2694 (1995). doi:10.1039/P19950002685
14. Olgivie, K. K. & Entwistle, D. W. Isomerization of tert-butyldimethylsilyl protecting groups in ribonucleosides. *Carbohydr. Res.* **89,** 203–210 (1981).
15. Mukaiyama, T., Murai, Y. & Shoda, S. An efficient method for glucosylation of hydroxy compounds using glucopyranosyl fluoride. *Chem. Lett.* **10,** 431–432 (1981).
16. Mukaiyama, T., Hashimoto, Y. & Shoda, S.-I. Stereoselective synthesis of 1,2-cis-glycofuranosides using glycofuranosyl fluorides. **12,** *Chem. Lett.* 935–938 (1983).
17. Paquette, L. A. & Bailey, S. Evaluation of D-Ribose as an Enantiopure Building Block for Construction of the C-Ring of Taxol and Its Congeners. *J. Org. Chem.* **60,** 7849–7856 (1995).
18. Lear, M. J., Yoshimura, F. & Hirama, M. A Direct and Efficient ―-Selective Glycosylation Protocol for the Kedarcidin Sugar, l-Mycarose: $AgPF_6$ as a Remarkable Activator of 2-Deoxythioglycosides. *Angew. Chem. Int. Ed.* **40,** 946–949 (2001).
19. Mikhailov, S. N., Kulikova, I. V., Nauwelaerts, K. & Herdewijn, P. Synthesis of 2′-O-α-d-ribofuranosyladenosine, monomeric unit of poly(ADP–ribose). *Tetrahedron* **64,** 2871–2876 (2008).
20. van der Heden van Noort, G. J., Overkleeft, H. S., van der Marel, G. A. & Filippov, D. V. Ribosylation of Adenosine: An Orthogonally Protected Building Block for the Synthesis of ADP-Ribosyl Oligomers. *Org. Lett.* **13,** 2920–2923 (2011).
21. Larsen, C. H., Ridgway, B. H., Shaw, J. T., Smith, D. M. & Woerpel, K. A. Stereoselective C-Glycosylation Reactions of Ribose Derivatives: Electronic Effects of Five-Membered Ring Oxocarbenium Ions. *J. Am. Chem. Soc.* **127,** 10879–10884 (2005).
22. van Rijssel, E. R. *et al.* Furanosyl Oxocarbenium Ion Stability and Stereoselectivity. *Angew. Chem. Int. Ed.* **53,** 10381–10385 (2014).
23. Aritomo, K., Wada, T. & Sekine, M. Alkylation of 6-N-acylated adenosine derivatives by the use of phase transfer catalysis. *J. Chem. Soc., Perkin Trans. 1* 1837 (1995).
24. Zhu, X.-F., Williams, H. J. & Ian Scott, A. Aqueous Trichloroacetic Acid: Another Useful

Reagent for Highly Selective 5′-Desilylation of Multisilylated Nucleosides. *Synth. Commun.* **33,** 2011–2016 (2003).
25. Ireland, R. E. *et al.* The Total Synthesis of Ionophore Antibiotics. A Convergent Synthesis of Lasalocid A (X537A). *J. Am. Chem. Soc.* **105,** 1988–2006 (1983).
26. Taylor, C. M., Barker, W. D., Weir, C. A. & Park, J. H. Toward a General Strategy for the Synthesis of 3,4-Dihydroxyprolines from Pentose Sugars. *J. Org. Chem.* **67,** 4466–4474 (2002).
27. Kistemaker, H. A. V. *et al.* Synthesis of well-defined adenosine diphosphate ribose oligomers. *Angew. Chem. Int. Ed. Engl.* **54,** 4915–4918 (2015).
28. Malpass, J. R., Hemmings, D. A., Wallis, A. L., Fletcher, S. R. & Patel, S. Synthesis and nicotinic acetylcholine-binding properties of epibatidine homologues: homoepibatidine and dihomoepibatidine. *J. Chem. Soc., Perkin Trans. 1* 1044–1050 (2001).
29. Fulmer, G. R. *et al.* NMR Chemical Shifts of Trace Impurities: Common Laboratory Solvents, Organics, and Gases in Deuterated Solvents Relevant to the Organometallic Chemist. *Organometallics* **29,** 2176–2179 (2010).

Chapter 3. Biological Studies Using the ADP-ribose Dimer and Fragments

Portions of this chapter are reprinted with permission from: Lambrecht, M. J.; Brichacek, M.; Barkauskaite, E.; Ariza, A.; Ahel, I.; and Hergenrother, P. J. *J. Am. Chem. Soc.* **2015**, *137*, 3558-3564. This work was performed in close collaboration with Dr. Matthew Brichacek. X-ray crystallography was performed by Dr. Eva Barkuskaite, Dr. Antonio Ariza and Prof. Ivan Ahel at the University of Oxford. ADP-HPD and ADP-HPM were synthesized by Dr. Bryan Yestrepsky.

3.1 Introduction

With the synthesis of the ADP-ribose dimer (**2-1**) complete, we sought to demonstrate its ability to enable biological experiments that were otherwise challenging or impossible to perform. For example, at the time of the work herein only one x-ray co-crystal structure of a PAR oligomer in complex with a protein had been published[1]. The lack of such examples was mostly due to the challenges associated with accessing homogeneous PAR oligomers and polymers as detailed in chapter 1. Additionally, analysis of PAR-protein binding in a quantitative manner remained difficult due to the heterogeneity of PAR. Prior to this work, most methods to characterize PAR-protein binding interactions were inconvenient and not amenable to high-throughput screening. Finally, while the ADP-ribose dimer (**2-1**) had been fully characterized by multidimensional NMR upon its synthesis, the lack of a literature spectrum with which to compare to complicated its analysis. Therefore, the work herein served to further prove its structure by demonstrating its competence in biological experiments.

3.2 The Processing of the ADP-ribose Dimer (**2-1**) by PARG and ARH3

Having accessed the ADP-ribose dimer (**2-1**), it was first important to assess the ability of PARG to hydrolyze it to two units of ADP-ribose. This experiment would both serve as further conformation of the structure of the ADP-ribose dimer (**2-1**) and would also show that it is a biochemically competent substrate for PARG. Interestingly, the ADP-ribose dimer (**2-1**) had never been explicitly shown to be a substrate for PARG, but previous work by the Ahel lab suggested that it would be[1,2]. When comparing PAR polymers treated with F398G *Tetrahymena thermophilia* PARG or F902G human PARG (inactive mutants) to the appropriate wild type enzymes, it was found that the amount of ADP-ribose dimer (**2-1**) detected by LCMS/MS analysis for the wild type enzymes was substantially decreased relative to treatment with the mutants[1]. However, substantial amounts of the ADP-ribose dimer (**2-1**) did still remain upon completion of wild type treatment in this assay (and to a lesser extent the trimer and tetramer), especially for human PARG[1]. It was therefore not clear when these studies commenced how good of a substrate this dimer was for PARG, especially for the human enzyme.

To address this question, the ADP-ribose dimer (**2-1**) was treated with PARG from various sources and conversion was monitored by HPLC (Figure 3.1). PARG from *Bos taurus* (bovine), *Tetrahymena thermophilia* (protozoan), and *Thermomonospora curvata*[3] (bacterial) was found to catalyze the hydrolysis of the ADP-ribose dimer (**2-1**) to two molecules of ADP-ribose. ARH3, another enzyme with PAR processing activity[4] was also found to possess activity against the dimer. The necessity of high enzyme concentrations and long reaction times are consistent with previous reports of the reduced *in vitro* activity of ARH3 when compared to PARG[5]. Treatment of the ADP-ribose dimer (**2-1**) with human PARG was evaluated by an LCMS assay (Figure 3.2), and it was confirmed that the product of the reaction was ADP-ribose

by MS analysis (the appearance of ADP-ribose in the HPLC assay was confirmed by retention time comparison to standards but not by MS). Importantly, catalytically inactive *T. thermophilia* (E256Q, Figure 3.1a) and human (E755N and E756N, Figure 3.2a) PARGs showed no processing of the ADP-ribose dimer (**2-1**) even at long times and high concentrations of enzyme.

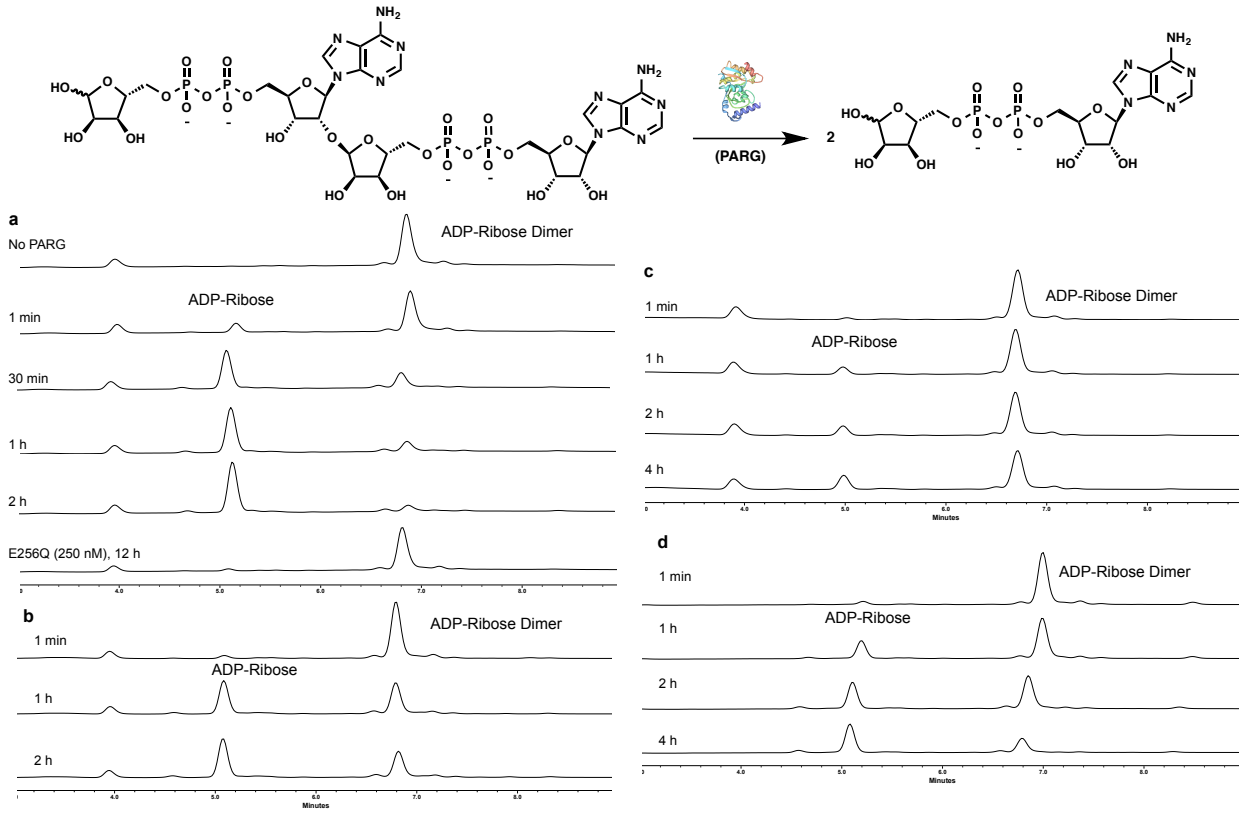


Figure 3.1. Processing of ADP-ribose dimer (**2-1**) by PARG from various sources as analyzed by HPLC. a) 25 nM *T. thermophilia* PARG without enzyme, with enzyme after 1 min, 30 min, 1 h, 2 h and with 250 nM E256Q (mutant) at 12 h. b) 5 nM B. taurus PARG at 1 min, 1 h, and 2 h. c) 250 nM *T. curvata* PARG at 1 min, 1 h, 2 h, and 4 h and d) 250 nM human ARH3 at 1 min, 1 h, 2 h, and 4 h. The peak near 4 min in panels a, b, and c is found in all traces containing the PARG assay buffer.

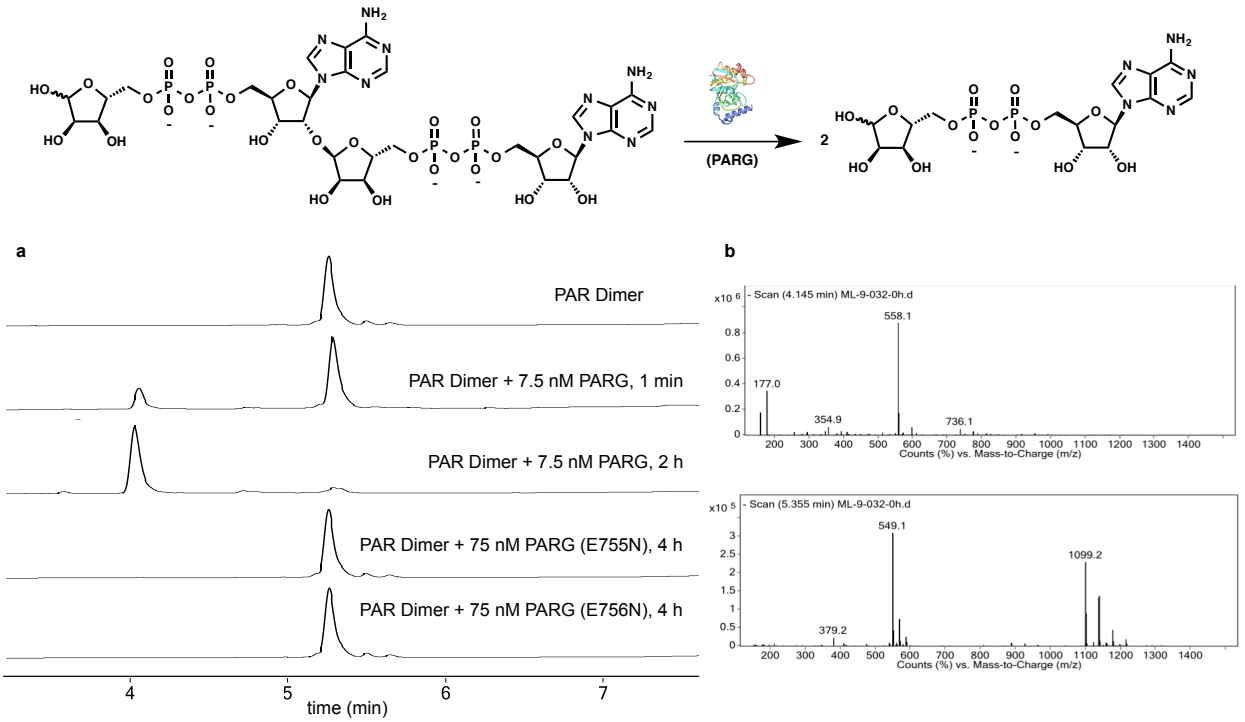


Figure 3.2. Processing of ADP-ribose dimer (**2-1**) by Human PARG as analyzed by LCMS. a) Treatment of ADP-ribose dimer (**2-1**) with 7.5 nM PARG at 1 min, 2 h and treatment with 75 nM E755N and E756N mutant PARGs at 4 h. b) Mass spectra at 4.15 min, 5.36 min for traces in a. ADP-Ribose ($[M-H]^-$ = 558.1) is shown on top, while the ADP-ribose dimer (**2-1**) ($[M-H]^-$ = 1099.2 and $[M-2H]^{2-}$ = 549.1) is shown on bottom.

3.3 ADP-ribose Dimer (**2-1**)/Human PARG X-Ray Structure

The inability of mutant PARGs to processes the ADP-ribose dimer (**2-1**) offered the opportunity to solve a co-crystal structure of the ADP-ribose dimer (**2-1**) in complex with human PARG. Though X-ray structures of human PARG existed[6], including those bound to the inhibitors ADP-HPD and ADP-ribose, such structures did not exist with human PARG bound to its natural substrate. Through a collaboration with Prof. Ivan Ahel's lab, a co-crystal structure of the ADP-ribose dimer (**2-1**) was solved in complex with human E756N PARG to 1.9 Å (Figure 3.3). Such a structure could not be solved with E755N human PARG, as only monomeric ADP-ribose was seen upon crystallization. This result suggests that the E755N mutant retains small

amounts of activity, especially at the high ratios of enzyme to substrate in crystallography experiments.

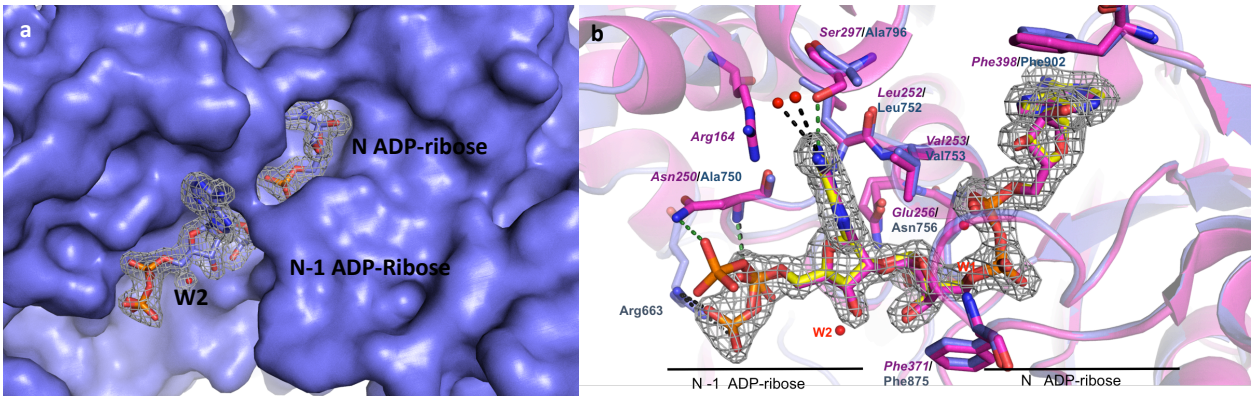


Figure 3.3. Human PARG-ADP-ribose dimer co-crystal structure. a) Surface representation of human PARG (E756N) bound to the ADP-ribose dimer (**2-1**). The gray mesh shows the $2F_O$-F_C electron density corresponding to the ordered region of the ADP-ribose dimer (**2-1**) at a counter level of 1σ. b) Overlay of protozoan (PDB:4L2H) and human PARG (this work) structures bound to a PAR fragment (pink) and the ADP-ribose dimer (**2-1**) (yellow), respectively. Figures by Dr. Eva Barkuskaite, Dr. Antonio Ariza, and Prof. Ivan Ahel (University of Oxford).

Examination of the human PARG/ADP-ribose dimer co-crystal structure reveals many similarities to the previously solved *T. thermophilia* PARG structure with short oligomers obtained through enzymatic synthesis and fractionation,[1] but also some differences. As with the *T. thermophilia* structure, binding appears to be tightest to the N ADP-ribose unit (Figure 3.3a). Most interactions with the N ADP-ribose unit are maintained in both structures, and the polymer is in a nearly identical conformation in both (Figure 3.3b). The human PARG structure offers the first view of the glycosidic linkage and N-1 ADP-ribose unit for the human protein and some differences with respect to *T. thermophilia* exist. Specifically, human PARG lacks a residue homologous to Arg164, and therefore the N-1 adenine in the human system is more exposed. The analogous residue to *T. thermophilia* Ser297 in the human protein is Ala796, and this replacement lacks contacts with the N-1 adenine again leaving it further solvent exposed. The N-1 β phosphate in the human protein forms hydrogen bonds with Arg663 which are not present in

the protozoan structure. The N-1 ribose residue is not seen in either structure due to disorder in this region.

3.4 Interrogation of the Minimal PARG Substrate

In an attempt to understand what portions of the PAR polymer are recognized and processed by PARG, the synthesis of a series of analogs and fragments of PAR was undertaken. This work was done with the goal of finding simplified analogs that PARG might bind to and process from which inhibitors and substrates could be designed. Specifically, if the presence of the pyrophosphate motif was not required for PARG substrate recognition and processing, inhibitors that lack this motif could be synthesized and would likely have better cell permeability properties than ADP-HPD (the best current *in vitro* inhibitor of PARG). Discovery of simplified substrates could also enable design of molecules that could be used in activity assays for PARG for applications such as screening.

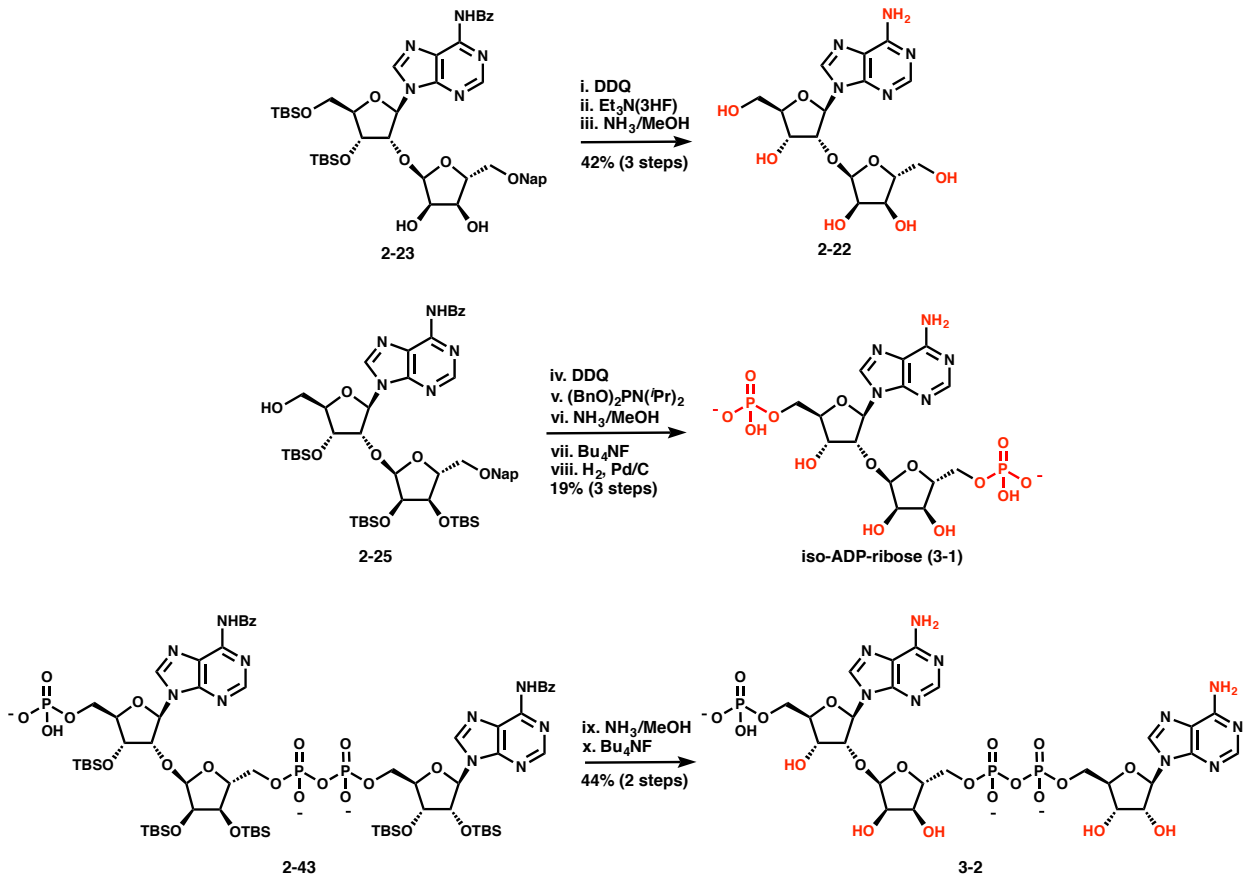


Figure 3.4. Synthesis of PAR fragments. i. DDQ (1.1 eq), CH_2Cl_2/H_2O, 0°C, 5 h. ii. $Et_3N{\cdot}3HF$, CH_2Cl_2, rt, 5 h, iii. NH_3/MeOH, rt, 12 h, 42% (for 3 steps). iv. DDQ (1.5 eq), CH_2Cl_2/H_2O, 0°C, 2 h, 79%. v. $(BnO)_2PN(^iPr)_2$ (2.5 eq), dicyanoimidazole (2.5 eq), CH_3CN, rt, 1 h, then: tBuOOH (10 eq), rt, 1 h, 70%. vi. NH_3/MeOH, rt, 15 h. vii. Bu_4NF, THF, rt, 3 h. viii. H_2, Pd/C, Et_3N (6 eq), $tBuOH/H_2O$, 16 h, ion pairing chromatography, cation exchange, 19% (3 steps). ix. NH3/MeOH, rt, 15 h. x. Bu_4NF, THF, rt, 3.5 h, 44% (2 steps).

Using intermediates of the synthetic route described in chapter 2, PAR fragments were synthesized (Figure 3.4). The central disaccharide of PAR (**2-22**) had been previously synthesized to confirm the structure of the glycosylated product (section 2.1.4). Alcohol **2-25** was used as a starting material to synthesize iso-ADP-ribose (**3-1**). Treatment of alcohol **2-25** with DDQ afforded a diol that after a bis phosphoramidite coupling, oxidation and global deprotection afforded iso-ADP-ribose (**3-1**). While iso-ADP-ribose has been synthesized enzymatically by digestion of PAR with snake venom phosphodiesterase, this represented the first chemical synthesis of this widely used molecule[7-9]. Global deprotection of phosphate **2-43**

afforded compound **3-2**, an analog very similar to the portion of the ADP-ribose dimer (**2-1**) that is visible in the PARG x-ray structure (Figure 3.3), lacking only the N-1 β phosphate.

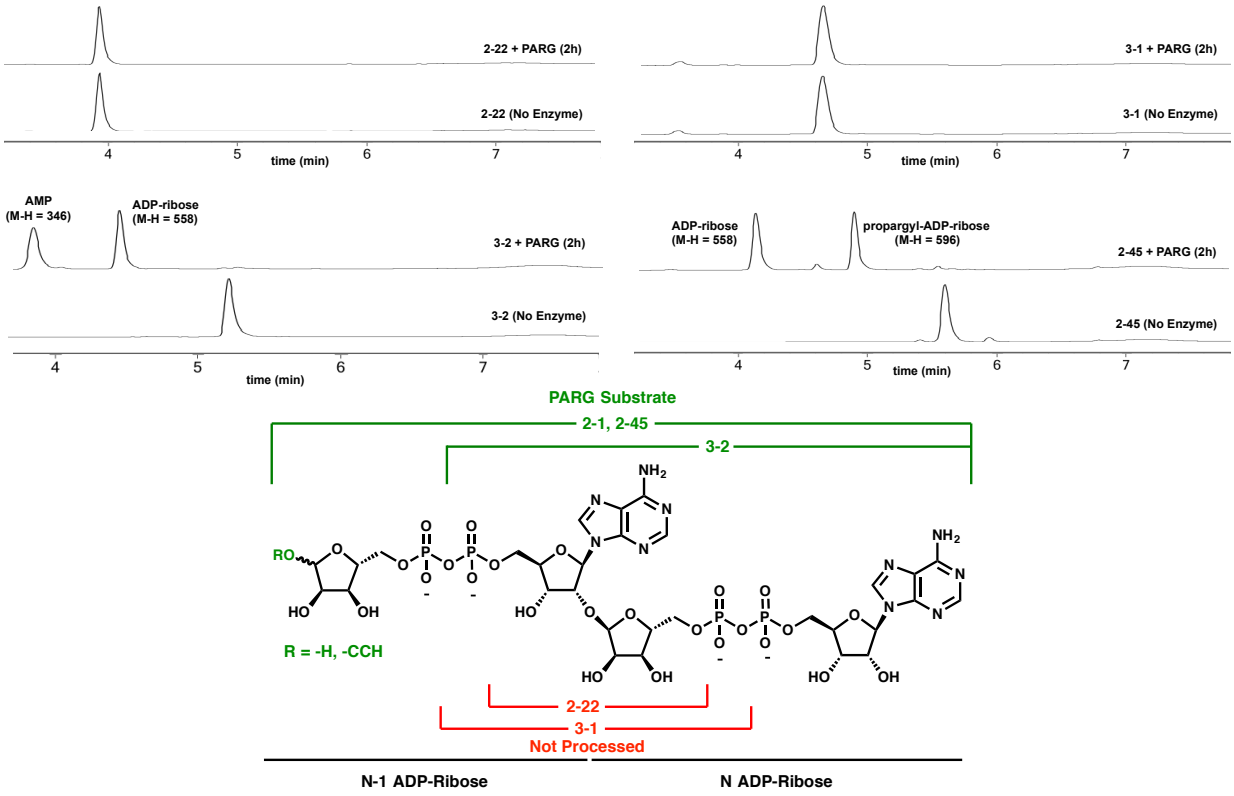


Figure 3.5. Processing of ADP-ribose dimer derivatives **2-22**, **2-45**, **3-1**, and **3-2** by *T. thermophilia* PARG as analyzed by LCMS. In addition to dimer **2-1**, compounds **2-45** and **3-2** retain the ability to be processed by PARG in this assay. Compounds **2-22** and **3-1** are not processed under the assay conditions.

The ability of these compounds to be processed by PARG was next examined (Figure 3.5). Neither disaccharide **2-22** nor iso-ADP-ribose (**3-1**) was found to be a substrate for PARG, even at high concentrations of enzyme. Phosphate **3-2** and the propargyl ADP-ribose dimer (**2-45**) were both found to be good substrates for the enzyme, as the glycosidic linkage was hydrolyzed with similar efficiency to the natural ADP-ribose dimer (**2-1**) under the conditions tested. This result was important in that it suggests that at least for PARG, the alkyne substitution to the N-1 ADP-ribose is not deleterious to the properties of the molecule. Therefore, the biotin and fluorophore containing probe molecules (synthesized in chapter 2) are likely to retain

biological activity. Also, because the propargyl ADP-ribose dimer (**2-45**) contains the unnatural β-stereochemical linkage at the N-1 ribose, the enzyme does not process this molecule to two molecules of ADP-ribose, but rather one molecule of ADP-ribose and one molecule of propargyl ADP-ribose.

3.5 Development of a PAR-protein Binding Assay

The ability of ADP-ribose oligomers and polymers to bind proteins is known, however quantifying such interactions is often challenging. Additionally few methods exist for high throughput screening for compounds that inhibit or disrupt PAR-protein binding. It was imagined that the Cy3 ADP-ribose dimer (**2-45**) may present a potential solution to these problems could its binding to proteins be detected by fluorescence polarization. A fluorescence polarization-based binding assay was therefore developed (Figure 3.6) using this molecule.

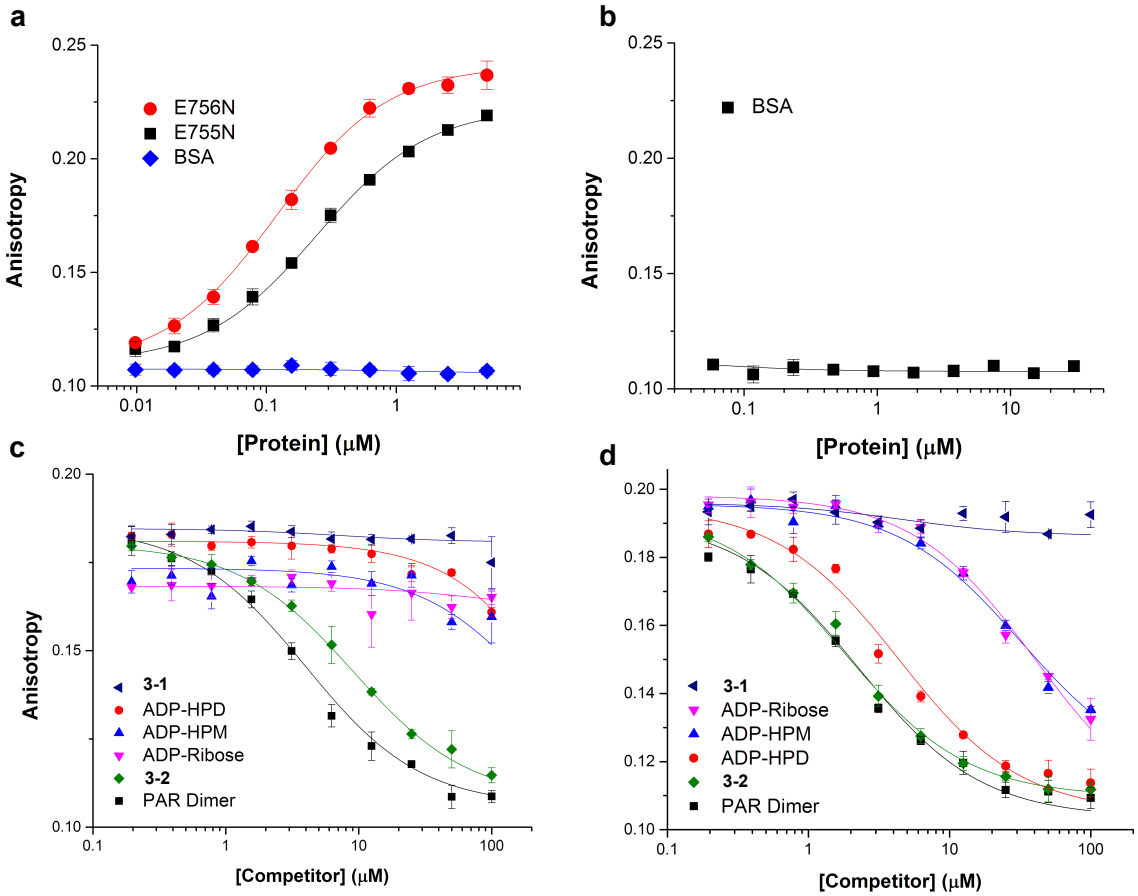


Figure 3.6. PAR-protein binding assay by fluorescence polarization. a) The Cy3 ADP-ribose dimer (**2-53**) binds to E755N (red circles) and E756N (black squares) PARG, but not to bovine serum albumin (BSA, blue diamonds). b) The Cy3 ADP-ribose dimer (**2-53**) does not bind to BSA even at high concentrations of protein. c) Fragments and inhibitors compete for binding with the Cy3 ADP-ribose dimer (**2-53**) to E755N PARG. d) Fragments and inhibitors compete for binding with the Cy3 ADP-ribose dimer (**2-53**) to E756N PARG.

The Cy3 ADP-ribose dimer (**2-53**) exhibited good binding to both E755N and E756N human PARG (Figure 3.6a), with K_d values of 208 ± 14 and 83 ± 7 nM, respectively. Interaction between the Cy3 ADP-ribose dimer (**2-53**) and bovine serum albumin (BSA), a protein that does not bind PAR, was not seen (Figure 3.6a) even at high concentrations of the protein (Figure 3.6b). Binding of the Cy3 ADP-ribose dimer (**2-53**) to the PARG mutants could be competed away with unlabeled ADP-ribose dimer (**2-1**) as well as with various compounds that inhibit

PARG (Figure 3.6 c and d), suggesting that this may be a viable assay for high-throughput screening of PAR-protein binding inhibitors. Interestingly, the ability of ADP-HPD, ADP-HPM and ADP-ribose to compete for binding with the Cy3 ADP-ribose dimer (**2-53**) is much reduced for E755N relative to E756N PARG (Figure 3.6c vs. 3.6d). The exact origins of this discrepancy are not yet well understood. The traditional role of E755 is believed to be a hydrogen bonding interaction with the 2''-OH of the N ADP-ribose unit and in most cases, mutation of this residue leads to abolishment of glycohydrolase activity. Therefore, it could be that mutation of E755 leads to reduced binding of the PAR polymer (and by extension inhibitors) to PARG. However, the Cy3 ADP-ribose dimer (**2-53**) does still bind well to this protein (only a 2 fold reduction in K_d) and the ADP-ribose dimer (**2-1**) and compound **3-2** do still compete away binding. Attempts by Dr. Bryan Yestrepsky and Tomohiro Shiari to strengthen the interaction of ADP-HPD to this residue in the wild type protein by installation of an amine in the 2'' position of APD-HPD have been unsuccessful.

3.6 Conclusions

The results presented here show the ability of the chemical synthesis of dimeric ADP-ribose and the synthetic route designed in chapter 2 to impact understanding of PAR biology. The chemical synthesis of dimeric ADP-ribose (**2-1**) enabled the first x-ray co-crystal structure of human PARG in complex with a natural substrate. Additionally, a brief SAR of PARG binding and processing of fragments was undertaken using molecules that could be easily synthesized from intermediates on the route to dimeric ADP-ribose (**2-1**). While these studies in many ways show why PARG inhibitors such as ADP-HPD and ADP-HPM are good inhibitors (strong binding of PARG to ADP-ribose motifs), they fail to inform design of better inhibitors of PARG that do not contain the pyrophosphate motif that limits cell permeability.

3.7 Experimental

Crystallographic Data and Model Refinement Parameters

Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, b, c (Å)	67.5, 91.3, 96.12
α, β, γ (°)	90, 90, 90
Resolution	55.22-1.95 (2.00-1.95)
R_{merge} (%)	10.5 (58.1)
I/sigma	21.4 (4.9)
Completeness	99.9 (99.8)
Multiplicity	13.2 (12.1)
Refinement	
Resolution	55.22-1.95
No. unique reflections	44991
R_{work}/R_{free}	13.7 / 17.9
No of atoms	
Protein	4202
Ligand/ion	205
Water	421
r.m.s. deviations	
Bond lengths (Å)	0.019
Bond angles (°)	1.55
B factors (Å^2)	
Protein	23.025
Ligand/ion	43.813
Water	35.080

Methods

Plasmids and proteins

T. thermophila PARG (TTHERM_00294690) and human PARG (a.a. 448–976 [K616A, Q617A, K618A, E688A, K689A, K690A]) proteins were expressed and purified as described previously[1,6] with the following changes. 6His-TEV-tagged catalytic mutant human PARG was obtained by introducing the E756N mutation using the QuickChange II site-directed mutagenesis kit (Stratagene). The human PARG proteins were purified from the clarified lysate by

immobilised metal affinity chromatography (IMAC) using a 5 mL His-TRAP HP column (GE Healthcare). Pooled fractions enriched for 6His-TEV-PARG were incubated with 6-His-tagged TEV protease whilst being dialysed for 24 hours at 4ºC. Cleaved protein was separated from uncleaved material, free 6-His tags and TEV protease by subtractive IMAC. The cleaved protein was concentrated before loading on a Superdex 200 pg (16/600) column (GE Healthcare) for size exclusion chromatography. Pooled fractions containing human PARG were concentrated to 7.5 mg/mL for crystallisation studies. IMAC and size exclusion chromatography were performed on an ÄKTA PURE FPLC system (GE Healthcare).

Crystallization and structure solution

Crystals were grown at 293K by sitting-drop vapour diffusion by mixing 350 nL purified protein (at 7.5 mg/mL in 50mM HEPES, pH 7.0, 150 mM NaCl, 2mM DTT) briefly pre-incubated with 1mM of **1** dimer with 150 nL of seed stock and 500 nL of mother liquor consisting of 18-23% PEG-3350, 0.2 M ammonium sulphate, 0.1 M PCTP pH 7.5. Seed stock was prepared using a Seed Bead^{TM} (Hampton Research) from a co-crystal of mutant human PARG with ADP-ribose. The co-crystal mother liquor (19% PEG-3350, 0.2 M ammonium sulphate, 0.1 M PCTP pH 7.5) was used as the stabilising solution for the seed stock. Crystals appeared over 48 hours and continued to grow for a further week and were then cryoprotected in a solution of 20% glycerol in mother liquor prior to being flash-cooled in liquid nitrogen. The diffraction data were collected at the Diamond Light Source (UK), with the automatic processing carried out by *Xia2*. The structure was solved by molecular replacement using using *Phaser* and the previously reported human PARG structure, by iterative rounds of manual building in *Coot* and crystallographic refinement in *REFMAC5*. Crystallographic figures were generated using PyMOL (http://www.pymol.org).

PARG Catalyzed Degradation LC/MS Assay (Figure 3.2)

Concentrated PARG stock solutions were diluted to the appropriate concentration in PARG activity buffer (50 mM K_2HPO_4, 50 mM KCl, 9 mM β-mercaptoethanol, pH = 7.5). Next, to a 0.5 mL tube was added 80 mL of PARG/ARH3 activity buffer, 10 mL of a 1 mM solution of synthetic ADP-Ribose Dimer, and 10 mL of the diluted PARG/ARH3. The tube was vortexed and transferred to a 384-well plate. The sample was analyzed by LC/MS (10 mL injections) at the appropriate time points. The HPLC method (98% 5 mM pentylamine:HOAc (pH=6.5) with 2% CH_3CN to 75% 6 mM pentylamine:HOAc (pH=6.5) with 25% CH_3CN over 5 min, then hold 75% 6 mM pentylamine:HOAc (pH=6.5) with 25% CH_3CN for 3 min) results in elution of ADP-Ribose at 4 min and PAR dimer at 5.2 min.

PARG or ARH3 Catalyzed Degradation HPLC Assay (Figure 3.1)

Concentrated PARG/ARH3 stock solutions were diluted to the appropriate concentration in PARG activity buffer (50 mM K_2HPO_4, 50 mM KCl, 9 mM β-mercaptoethanol, pH = 7.5) or ARH3 activity buffer (50 mM PBS, 10 mM $MgCl_2$, 5 mM TCEP, pH = 7.1). Next, to a 0.5 mL tube was added 42.5 mL of PARG/ARH3 activity buffer, 5 mL of a 1 mM solution of synthetic ADP-Ribose Dimer, and 2.5 mL of the diluted PARG/ARH3. The tube was vortexed and transferred to an HPLC vial with a small volume insert. The sample was analyzed by HPLC (10 mL injections) at the appropriate time points. The HPLC method (99% 8 mM Et_3NHOAc/1% CH_3CN to 90% 8 mM Et_3NHOAc/10% CH_3CN) results in elution of ADP-Ribose at 5.1 min and PAR dimer at 6.9 min.

Fluorescence Polarization Assays

E755N and E756N PARG dilutions (final concentrations of 5 μM to 0 μM) were made in PARG binding buffer (50 mM KCl, 50 mM K_2HPO_4, 0.05% Trition X-100, pH = 7.5) and

compound **22** was added to a final concentration of 7.5 nM. The solutions were transferred to a black 384 well plate incubated at room temperature for 30 min and read using an Analyst HT (excitation at 530 ± 25 nm and emission at 570 ± 10 nm). Polarization data was converted to anisotropy, plotted and fit to the following equation[10,11] using Origin Lab:

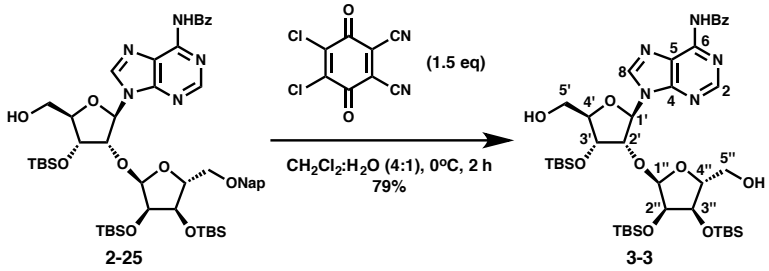
$$A_{OBS} = A_F + \frac{(A_B - A_F)\left(K_D + L_{ST} + R_T - \sqrt{((K_D + L_{ST} + R_T)^2 + 4L_{ST}R_T)}\right)}{2L_{ST}}$$

where A_{OBS} is the observed anisotropy, A_F is the anisotropy of the free probe, A_B is the anisotropy of the bound probe K_D is the dissociation constant of the protein-ligand (fluorescently labeled compound) interaction, L_{ST} is the total concentration of the labeled ligand in the assay, and R_T is the total receptor (protein) concentration. The error of the reported K_D values corresponds to the error in the curve fitting.

For competition binding experiments, unlabeled compound **1** or ADP-HPD was diluted to the appropriate concentrations and protein in binding buffer was added to final concentrations for E755N and E756N of 500 nM and 250 nM, respectively. The solutions were incubated for 30 min at room temperature at which point compound **22** was added at a final concentration of 7.5 nM. The solutions were transferred to a 384 well black plate incubated an additional 30 min and read using the conditions described above. Polarization data was converted to anisotropy, plotted and fit to the following equation using Origin Lab:

$$A = A_F + \frac{A_B - A_F}{1 + 10^{log(L_T - IC_{50})}}$$

where L_T is the concentration of unlabeled competitor and IC_{50} is the concentration of unlabeled competitor necessary to produce 50% decrease in binding of the labeled probe.

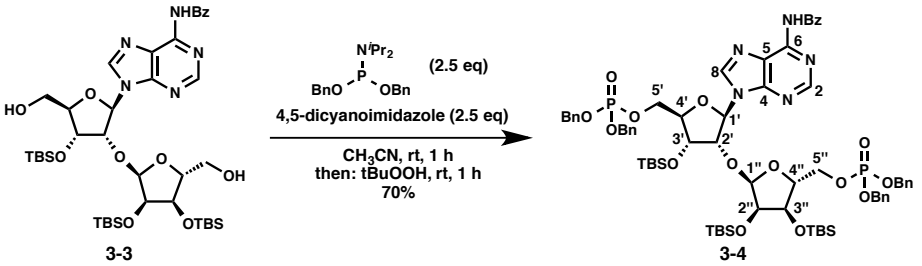


Compound 3-3: To a 0°C solution of compound **2-25** (50 mg, 0.051 mmol, 1 eq) in CH_2Cl_2 (0.9 mL) and H_2O (0.2 mL) was added 2,3-Dichloro-5,6-dicyano-1,4-benzoquinone (17.3 mg, 0.076 mmol, 1.5 eq). The solution was stirred at 0°C for 2 h, poured into saturated aqueous $Na_2S_2O_3$, extracted three times with ethyl acetate, dried through Na_2SO_4, concentrated and purified by silica column chromatography to afford compound **3-3** (34 mg, 79%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 9.19 (s, 1H, -N*H*Bz), 8.79 (s, 1H, H-8), 8.09 (s, 1H, H-2), 8.02 (d, J = 7.7 Hz, 2H, -Bz), 7.62 (t, J = 7.5 Hz, 1H, -Bz), 7.53 (t, J = 7.7 Hz, 2H, -Bz), 6.07 (d, J = 7.6 Hz, 1H, H-1'), 5.02 (dd, J = 7.7, 4.3 Hz, 1H, H-2'), 4.82 (d, J = 3.1 Hz, 1H, H-1''), 4.59 (d, J = 4.3 Hz, 1H, H-3'), 4.21 (s, 1H, H-4'), 4.10 (d, J = 6.7 Hz, 1H, H-3''), 4.02 – 3.91 (m, 2H, H-4'' + 5'a), 3.87 – 3.69 (m, 3H, 2'' + 5'b + 5''a), 3.53 (d, J = 12.1 Hz, 1H, H-5''b), 0.94 (s, 9H, -tBu), 0.88 (s, 9H, -tBu), 0.85 (s, 9H, -tBu), 0.16 (s, 3H, -CH_3), 0.13 (s, 3H, -CH_3), 0.03 (s, 3H, -CH_3), 0.01 (s, 3H, -CH_3), -0.00 (s, 3H, -CH_3), -0.05 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 164.7, 152.4, 150.8, 150.5, 143.9, 133.7, 133.2, 129.2, 128.1, 124.7, 104.9, 89.7, 89.1, 82.8, 81.9, 73.9, 73.5, 70.6, 63.0, 60.9, 26.3, 26.2, 26.2, 18.8, 18.4, 18.4, -3.9, -4.0, -4.1, -4.4, -4.7.

HRMS (ESI) m/z calcd for $C_{40}H_{68}N_5O_9Si_3$ ($[M+H]^+$) 846.4325, found 846.4318.



Compound 3-4: To a dry 4 mL reaction vial was added compound **3-3** (19.0 mg, 0.0225 mmol, 1 eq) and dibenzyl N,N-diisopropylphosphoramidite (19.4 mg, 0.562 mmol, 2.5 eq). The compounds were co-evaporated three times on a N_2-filled rotary evaporator with dry CH_3CN and then dried over P_2O_5 overnight. Separately, 4,5-dicyanoimidazole (6.6 mg, 0.0562 mmol, 2.5 eq) was evaporated and dried in a similar manner. To compound **3-3** and the phosphoramidite was added dry CH_3CN (0.1 mL), followed by a solution of the 4,5-dicyanoimidazole in CH_3CN (0.1 mL) at room temperature. The reaction was stirred for 1 h at which point tBuOOH (0.041 mL of a 5.5 M solution, 0.225 mmol, 10 eq) was added and the reaction stirred an additional hour. The reaction was quenched with saturated aqueous $NaHCO_3$, extracted three times with ethyl acetate, dried through Na_2SO_4, and purified by silica column chromatography to give compound **3-4** (21.5 mg, 70%).

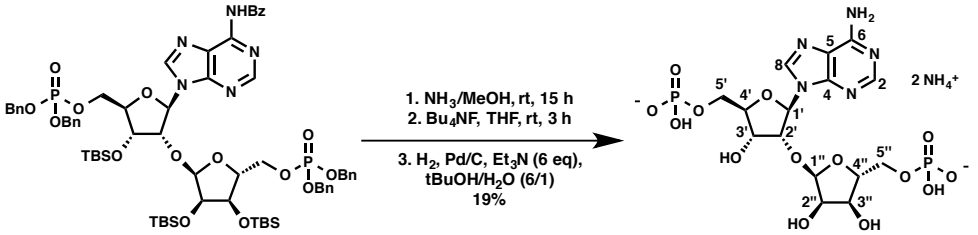
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.96 (bs, 1H, -N*H*Bz) 8.69 (s, 1H, H-8), 8.24 (s, 1H, H-2), 7.98 (d, *J* = 7.5 Hz, 2H, -Bz), 7.63 – 7.56 (m, 1H, -Bz), 7.51 (t, *J* = 7.7 Hz, 2H, -Bz), 7.37 – 7.25 (m, 20H, -Bn), 6.11 (d, *J* = 2.8 Hz, 1H, H-1'), 5.19 (d, *J* = 3.7 Hz, 1H, H-1''), 5.03 – 4.88 (m, 8H, -C*H*$_2$Ph), 4.71 (dd, *J* = 4.5, 2.9 Hz, 1H, H-2'), 4.46 (dd, *J* = 6.6, 4.5 Hz, 1H, H-3'), 4.35 (ddd, *J* = 11.4, 5.7, 3.4 Hz, 1H, H-5'a), 4.28 – 4.23 (m, 1H, H-4''), 4.18 (dd, *J* = 4.5, 2.7 Hz, 1H, H-4'), 4.15 – 4.05 (m, 2H, H-5'b, H-5''a), 3.97 (ddd, *J* = 11.2, 5.8, 3.6 Hz, 1H, H-5''b), 3.93 (appt, *J* = 4.9 Hz, 1H, H-3''), 3.89 (dd, *J* = 5.3, 3.7 Hz, 1H, H-2''), 0.84 (s, 18H, -tBu), 0.79 (s, 9H, -tBu),

0.05 (s, 3H, -CH_3), 0.03 (s, 3H, -CH_3), 0.00 (s, 3H, -CH_3), -0.02 (s, 3H, -CH_3), -0.04 (s, 3H, -CH_3), -0.06 (s, 3H, -CH_3).

^{13}C NMR (125 MHz, $CDCl_3$) δ 164.7, 152.8, 151.2, 149.6, 142.2, 136.01 – 135.87 (m), 135.80 (d, *J* = 6.7 Hz), 133.0, 129.1, 128.8, 128.1, 120.2, 103.5, 88.7, 82.87 (d, *J* = 8.0 Hz), 82.26 (d, *J* = 8.3 Hz), 80.7, 73.6, 71.0, 70.3, 69.79 – 69.55 (m), 69.7 – 69.4 (m), 66.8 (d, *J* = 5.4 Hz), 65.5 (d, *J* = 5.3 Hz), 26.2, 26.0, 18.5, 18.3, 18.3, -4.1, -4.3, -4.4, -4.4, -4.5, -4.8.

^{31}P NMR (202 MHz, $CDCl_3$) δ 0.17, 0.05

HRMS (ESI) m/z calcd for $C_{66}H_{95}N_5O_{15}Si_3P_2Na$ ($[M+Na]^+$) 1366.5506, found 1366.5494.



3-4 3-1

iso-ADP-ribose (3-1): To a 7 mL reaction vial containing compound **3-4** (24 mg, 0.0176 mmol, 1 eq) was added ammonia solution (0.5 mL of a 7M solution in MeOH) and the reaction was stirred at room temperature for 15 h. The reaction was evaporated and THF (0.2 mL) and tetrabutylammonium fluoride (0.11 mL of a 1M solution in THF, 0.11 mmol, 6 eq) was added and the reaction was stirred at room temperature for 3 h at which time the solution was concentrated and run through a short column of silica. The eluent was collected, concentrated and transferred to a 7 mL reaction vial. Palladium on carbon (10 mg) was added followed by tBuOH (0.2 mL), water (0.03 mL) and triethylamine (11 mg, 0.109 mmol, 6 eq). A balloon filled with hydrogen was fitted to the flask and stirred at room temperature for 16 h. The solution was filtered through celite, concentrated and purified by preparative HPLC (gradient of 99% 8 mM Et_3NHOAc/1% CH_3CN to 90% 8 mM Et_3NHOAc/10% CH_3CN). The fractions containing

product were combined, concentrated and exchanged for the NH_4^+ cation using a Dowex 50W-X8 cation exchange column to give compound **3-1** (2.0 mg, 19%).

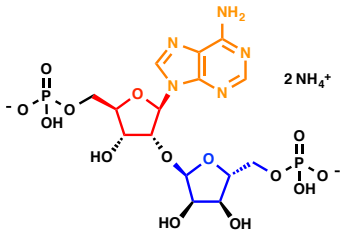
^{1}H NMR (500 MHz, D_2O) δ 8.57 (s, 1H, H-8), 8.27 (s, 1H, H-2), 6.29 (d, *J* = 5.6 Hz, 1H, H-1'), 5.22 (d, *J* = 4.1 Hz, 1H, H-1''), 4.86 (dd, *J* = 5.5, 5.2 Hz, 1H, H-2'), 4.62 (dd, *J* = 5.2, 3.6 Hz, 1H, H-3'), 4.45 – 4.39 (m, 1H, H-4'), 4.34 (appdt, *J* = 4.1, 2.2 Hz, 1H, H-4''), 4.18-4.16 (m, 2H, H-2'' + H-3''), 4.13 – 4.02 (m, 2H, H-5'a + H-5'b), 3.90 (ddd, *J* = 11.5, 5.9, 4.0 Hz, 1H, H-5''a), 3.84 (dt, *J* = 11.4, 4.3 Hz, 1H, H-5''b).

^{13}C NMR (125 MHz, D_2O) δ ^{13}C shifts determined from HSQC/HMBC, see table.

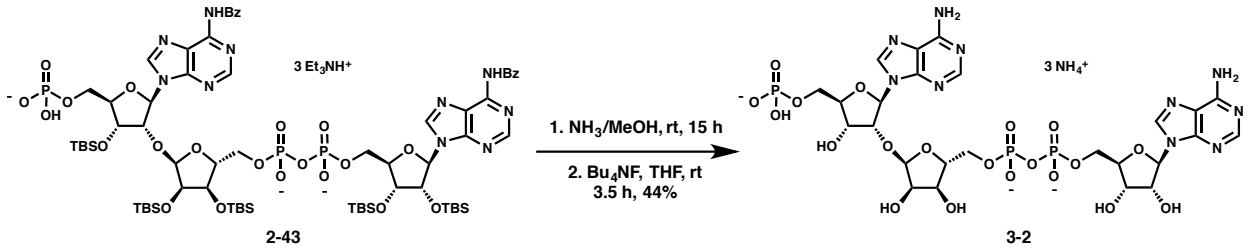
^{31}P NMR (242 MHz, D_2O) δ 2.12-1.92 (m, 2P)

HRMS (ESI) m/z calcd for $C_{15}H_{22}N_5O_{14}P_2$ ($[M-H]^-$) 558.0644, found 558.0637.

	^{1}H	^{13}C	^{31}P
Ado 1'	6.29 (d, *J* = 5.6 Hz)	86.1	
Ado 2'	4.86 (dd, *J* = 5.6, 5.2 Hz)	80.2	
Ado 3'	4.62 (dd, *J* = 4.9, 3.3 Hz)	70.8	
Ado 4'	4.42	84.9	
Ado 5'a	4.18	64.3	2.05
Ado 5'b	4.16		
Ade 2	8.27	152.6	
Ade 4	-	148.8	
Ade 5	-	118.5	
Ade 6	-	155.6	
Ade 8	8.57	140.3	
Rib 1''	5.22 (d, *J* = 4.1 Hz)	102.1	
Rib 2''	4.17	70.2	
Rib 3''	4.17	71.5	
Rib 4''	4.34	84.7	
Rib 5''a	3.90	64.5	2.05
Rib 5''b	3.84		



iso-ADP-ribose (3-1)



2-43 3-2

Compound 3-2: To a 7 mL vial containing compound **2-43** (5.7 mg, 2.9 mmol, 1 eq), was added NH_3/MeOH (0.5 mL of a 7M solution) and the reaction was stirred for 15 h. The solution was evaporated and Bu_4NF solution was added (0.2 mL of a 1 M solution in THF) and the reaction was stirred for 3.5 h. The reaction was quenched by addition of 3M aqueous NaOAc (0.2 mL) and stirred for 30 min. The solution was transferred to four 1.5 mL centrifuge tubes, EtOH (0.5 mL) was added to each tube, and the tubes were centrifuged for 5 min at 4°C and 14,000 rcf. The supernatant was discarded and the pellet was washed with EtOH (0.5 mL) and again centrifuged. The pellet was suspended in H_2O (1 mL) and purified by preparative HPLC (gradient of 99% 8 mM Et_3NHOAc/1% CH_3CN to 90% 8 mM Et_3NHOAc/10% CH_3CN). Compound **3-2** was obtained as the ammonium salt (1.1 mg, 44%) of the product was obtained after concentration and cation exchange (Dowex 50W-X8, NH_4^+ form).

^{1}H NMR (500 MHz, D_2O) δ 8.51 (s, 1H), 8.35 (s, 1H), 8.14 (s, 1H), 8.11 (s, 1H), 6.17 (d, *J* = 4.0 Hz, 1H), 6.00 (d, *J* = 5.9 Hz, 1H), 5.24 (d, *J* = 4.2 Hz, 1H), 4.69-4.64 (m, 2H), 4.58 (appt, *J* = 5.3 Hz, 1H), 4.47 (dd, *J* = 5.2, 3.5 Hz, 1H), 4.39 – 4.31 (m, 3H), 4.27 – 4.15 (m, 4H), 4.12 – 3.95 (m, 4H).

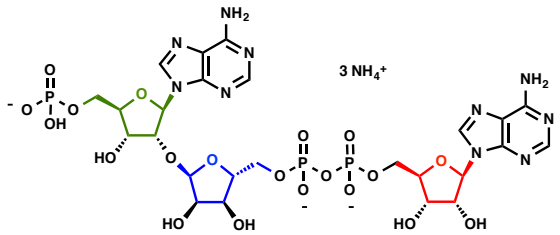
^{13}C NMR (125 MHz, D_2O) δ 155.6, 155.4, 153.0, 152.8, 148.9, 148.9, 140.1, 139.9, 118.6, 118.5, 101.7, 87.3, 86.6, 84.5, 84.5, 84.5, 79.7, 74.7, 74.6, 70.7, 70.2, 70.0, 65.8, 65.5, 64.9

Note: ^{13}C shifts determined from HSQC/HMBC

^{31}P NMR (242 MHz, D_2O) δ 3.71, -11.14 (d, *J* = 23.2 Hz), -11.27 (d, *J* = 22.1 Hz).

HRMS (ESI) m/z calcd for $C_{25}H_{35}N_{10}O_{20}P_3$ ([M-H]$^-$) 887.1169, found 887.1175.

	^{1}H	^{13}C	^{31}P
I-Ado 1'	6.00 (d, J = 5.9 Hz)	87.3	
I-Ado 2'	4.67	74.7	
I-Ado 3'	4.47 (dd, J =5.2, 3.5 Hz)	70.7	
I-Ado 4'	4.38	84.5	
I-Ado 5'a	4.22	65.5	-11.27
I-Ado 5'b	4.22		(d, J = 22.1 Hz)
α- Rib 1''	5.24 (d, J = 4.2 Hz)	101.7	
α- Rib 2''	4.20	70.2	
α– Rib 3''	4.20	71.6	
α- Rib 4''	4.38	84.5	
α- Rib 5''a	4.09	64.9	-11.14
α- Rib 5''b	4.03		(d, J = 23.2 Hz)
D-Ado 1'	6.17 (d, J = 4.0 Hz)	86.6	
D-Ado 2'	4.67	79.7	
D-Ado 3'	4.58 (appt, J = 5.3 Hz)	70.0	
D-Ado 4'	4.38	84.5	
D-Ado 5'a	4.09	65.8	3.71
D-Ado 5'b	4.03		



3-2

3.8 References

1. Barkauskaite, E. *et al.* Visualization of poly(ADP-ribose) bound to PARGreveals inherent balance between exo- andendo-glycohydrolase activities. *Nat. Commun.* **4,** 1–8 (2013).
2. Dunstan, M. S. *et al.* Structure and mechanism of a canonical poly(ADP-ribose) glycohydrolase. *Nat. Commun.* **3,** 878–6 (2012).
3. Slade, D. *et al.* The structure and catalytic mechanism of a poly(ADP-ribose) glycohydrolase. *Nature* **477,** 616–620 (2011).
4. Niere, M. *et al.* ADP-ribosylhydrolase 3 (ARH3), Not Poly(ADP-ribose) Glycohydrolase (PARG) Isoforms, Is Responsible for Degradation of Mitochondrial Matrix-associated Poly(ADP-ribose). *J. Biol. Chem.* **287,** 16088–16102 (2012).
5. Ono, T., Kasamatsu, A., Oka, S. & Moss, J. The 39-kDa poly(ADP-ribose) glycohydrolase ARH3 hydrolyzes O-acetyl-ADP-ribose, a product of the Sir2 family of

acetyl-histone deacetylases. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 16687–16691 (2006).
6. Tucker, J. A. *et al.* Structures of the Human Poly (ADP-Ribose) Glycohydrolase Catalytic Domain Confirm Catalytic Mechanism and Explain Inhibition by ADP-HPD Derivatives. *PLoS ONE* **7,** e50889 (2012).
7. Wang, Z. *et al.* Recognition of the iso-ADP-ribose moiety in poly(ADP-ribose) by WWE domains suggests a general mechanism for poly(ADP-ribosyl)ation-dependent ubiquitination. *Genes Dev.* **26,** 235–240 (2012).
8. Wang, Z., Gagné, J.-P., Poirier, G. G. & Xu, W. Crystallographic and Biochemical Analysis of the Mouse Poly(ADP-Ribose) Glycohydrolase. *PLoS ONE* **9,** e86010 (2014).
9. Zhang, F., Chen, Y., Li, M. & Yu, X. The oligonucleotide/oligosaccharide-binding fold motif is a poly(ADP-ribose)-binding domain that mediates DNA damage response. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 7278–7283 (2014).
10. Roehrl, M. H. A., Wang, J. Y. & Wagner, G. A General Framework for Development and Data Analysis of Competitive High-Throughput Screens for Small-Molecule Inhibitors of Protein−Protein Interactions by Fluorescence Polarization. *Biochemistry* **43,** 16056–16066 (2004).
11. Roehrl, M. H. A., Wang, J. Y. & Wagner, G. Discovery of Small-Molecule Inhibitors of the NFAT−Calcineurin Interaction by Competitive High-Throughput Fluorescence Polarization Screening. *Biochemistry* **43,** 16067–16075 (2004).

Chapter 4. Chemoenzymatic Synthesis of ADP-ribose Oligomers

Synthesis of the ADP-ribose dimer (**2-1**) as described in chapter 2 was performed in collaboration with Dr. Matthew Brichacek.

4.1 Introduction

The ability of the homogeneous ADP-ribose dimer (**2-1**) to enable otherwise challenging biological experiments as described in chapter 3 showed what a useful tool this compound is. More generally, these experiments showed the types of studies that could be performed with homogeneous ADP-ribose oligomers of defined chemical structure. It was imagined that could longer ADP-ribose oligomers and polymers be synthesized selectively and homogeneously, that they too could inform PAR biology. Unfortunately, the synthesis of such molecules by extension of the route detailed in chapter 2 was largely unsuccessful. The main obstacle inhibiting this route was the low yields and challenging purifications associated with solution phase pyrophosphate coupling reactions. One potential solution to this problem is the solid phase pyrophosphate formation that has been described recently[1,2]. However, though solid phase synthesis solves many of the problems associated with purification, yields for these processes are still generally low enough to not represent a general solution to PAR polymer synthesis. Therefore, we were inspired to adopt a different approach co-opting the chemical synthesis of substrates via controlled, enzymatic, PARP-enabled polymerization.

4.2 Approaches Towards Chemoenzymatic Synthesis of PAR Oligomers Using β-NAD^+ Analogs

As described in chapter 1, PARPs synthesize PAR from β-NAD^+ by extension of the polymer on the free 2' alcohol to form long and heterogeneous PAR chains (Figure 4.1a). Others have shown that unnatural β-NAD^+ analogs with deletion of the 2' alcohol can modify PARP, but terminate extension of PAR chains[3]. Therefore, we reasoned that ADP-ribose polymers and oligomers of homogenous length could be generated from β-NAD^+ analogs containing a removable protecting group (Figure 4.1b). It was envisioned that such oligomers would be incorporated by PARP and the excess substrate would be removed to afford homogeneously ADP-riboslyated proteins with a removable group on the 2' alcohol. Deprotection of this group and addition of more substrate would result in PARP homogeneously labeled with the ADP-ribose dimer. We envisioned this process being iterated through many rounds to afford PARP automodified with homogeneous PAR. Upon detachment of the polymers from the protein, homogeneous PAR could be obtained[4].

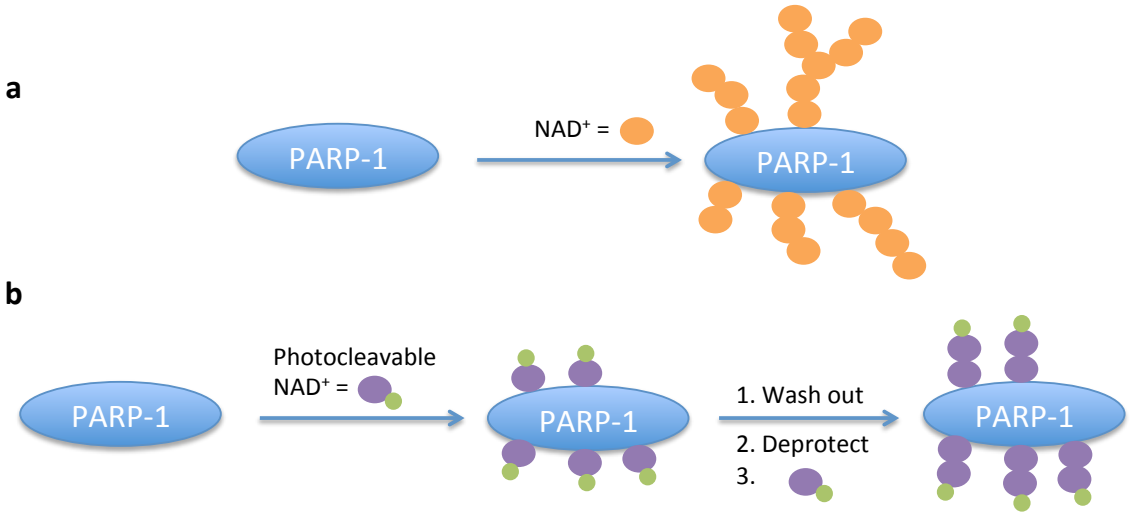


Figure 4.1. Synthesis of PAR oligomers by a) traditional methods and b) the envisioned chemoenzymatic approach with β-NAD^+ analogs.

4.2.1 Synthesis of β-NAD$^+$ Analogs for PARP Incorporation

It was imagined that a β-NAD$^+$ analog that could be iterated by PARP would need to possess several characteristics. First, perturbation of the β-NAD$^+$ analog from the natural substrate would need to be small enough that its incorporation by PARP was unaffected. Second, the protecting group would need to be removable under extremely mild conditions that are compatible with biological systems, as the growing polymer would still be protein bound when deprotected. Therefore, the 2-nitrobenzyl β-NAD$^+$ analog (**4-1**, Figure 4.2) was selected as an appropriate compound for these studies as the protecting group was small and its removal could be effected by treatment with UV light[5]. Additionally, the propargyl β-NAD$^+$ analog (**4-2**, Figure 4.2) was needed because it could be used to conjugate fluorescent tags, thereby enabling the monitoring of PARP incorporation of analogs. Oligomers and polymers synthesized containing the alkyne provided by **4-2** could also allow for modification of polymers in an analogous fashion to the propargyl APD-ribose dimer (**2-45**). Finally, a deoxy β-NAD$^+$ analog (**4-3**) was desired as it was similar to a compound that is a known substrate for PARP-1[3].

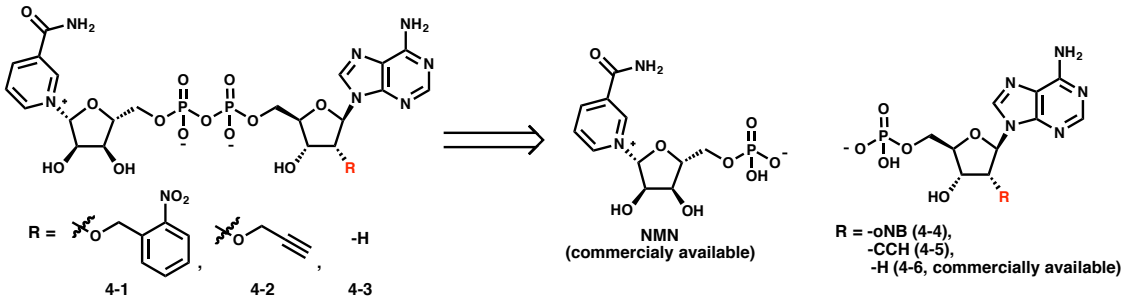


Figure 4.2. Retrosynthetic analysis of β-NAD$^+$ analogs **4-1**, **4-2**, and **4-3**.

As nicotinamide mononucleotide (NMN) is available commercially, it was envisioned that these analogs would originate from phosphorimidazolide couplings between NMN and

adenosine monophosphates with a 2-nitrobenzyl group (**4-4**), propargyl group (**4-5**), or hydrogen atom replacement (**4-6**) at the 2' alcohol. In contrast to the synthesis of the ADP-ribose dimer (**2-1**), the ultimate coupling in this case was performed without protection of the secondary alcohols due to the challenges of synthesis and deprotection of NMN[6] and its commercial availability in non-protected form.

In the forward sense (Figure 4.3), 2-nitrobenzyl adenosine monophosphate was synthesized by a known procedure involving treatment of adenosine with sodium hydride and 2-nitrobenzyl bromide to afford compound **4-7** as a single isomer after precipitation[7]. Similarly, propargylation with propargyl bromide and sodium hydride gave diol **4-8**[8]. Phosphorylation to afford phosphates **4-4** and **4-5** was performed using phosphorous oxychloride in trimethyl phosphate. Imidazolide formation followed by coupling with NMN gave 2-nitrobenzyl β-NAD^+ (**4-1**) and propargyl β-NAD^+ (**4-2**). Finally, deoxy β-NAD^+ was synthesized by imidazolide formation of commercially available deoxy adenosine monophosphate (**4-6**) and coupling to NMN to provide deoxy β-NAD^+ (**4-3**).

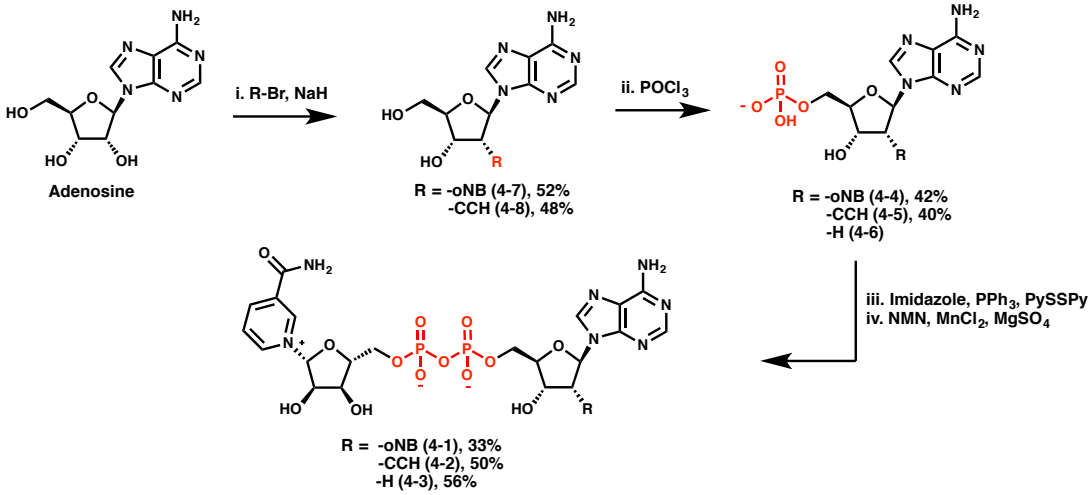


Figure 4.3. Synthesis of β-NAD^+ analogs **4-1**, **4-2** and **4-3**. i. R-Br (1.3-1.5 eq), NaH (1.1-1.3 eq), DMF, 0°C, 3 h, 52% for **4-7**, 48% for **4-8**. ii. $POCl_3$ (3 eq), $PO(OMe)_3$, 0°C, 1.5 h, reverse phase purification 42% for **4-4**, 40% for **4-5**. iii. Imidazole (16 eq), 2,2'-dipyridyldisulfide (3.2

eq), PPh_3 (3.2 eq), Et_3N (2.3 eq), DMSO, rt, 3 h. iv. NMN (0.8 eq), $MnCl_2$ (1.5 eq), $MgSO_4$ (2 eq), formamide, rt, 15 h, 33% for **4-1**, 50% for **4-2**, 56% for **4-3**.

4.2.2 Identification of an Active PARP for Chemoenzymatic Synthesis

With the appropriate NAD^+ analogs synthesized, an active enzyme capable of the synthesis of PAR oligomers was needed. Several enzymes possess PARP activity including PARPs-1, -2, and -3 as well as tankyrases-1 and -2 (PARPs -5a and -5b) [9]. Expression of full length PARP-1 is known to be challenging and while occasionally performed in *E. coli*[10], is most frequently performed in baculovirus systems in insect cells[11]. While expression of the PARP catalytic domain is known to proceed well in *E. coli*, its activity is reported to be substantially lower than the full length enzyme due to the requirement of the protein to bind DNA through its zinc finger domains[12]. The catalytic domain of tankyrase has been reported to have good, but reduced, PARP activity relative to full length PARP-1[13]. However, for some applications the ease of expression of catalytic tankyrase relative to PARP-1 has overcome its reduced activity[4]. We therefore decided to obtain and assess the ability of several enzymes (and domains of enzymes) to both synthesize PAR from β-NAD^+ as well as incorporate the unnatural analogs synthesized in chapter 4.2.1.

The catalytic domain of human tankyrase-1 was expressed in *E. coli* as a GST-fusion protein following a previously described procedure in our laboratory[14]. Several types of activity assays were performed to assess the ability of these proteins to incorporate natural β-NAD^+ as well as the unnatural analogs synthesized in chapter 4.2.1. Human tankyrase-1 was found to have low levels of PAR synthesis activity when analyzed by sypro ruby staining (Figure 4.4a). Some PAR formation by tankyrase-1 could be seen by immunoblot (Figure 4.4b), but incorporation of alkyne β-NAD^+ (**4-2**) was not observed by in-gel fluorescence after conjugation of Cy3 (Figure

4.4c). From these results we interpreted that while tankyrase-1 likely does produce PAR (as evident by immunoblot, figure 4b), such activity is likely low relative to full length PARP-1 and tankyrase-1 may not be able to incorporate unnatural β-NAD^+ analogs.

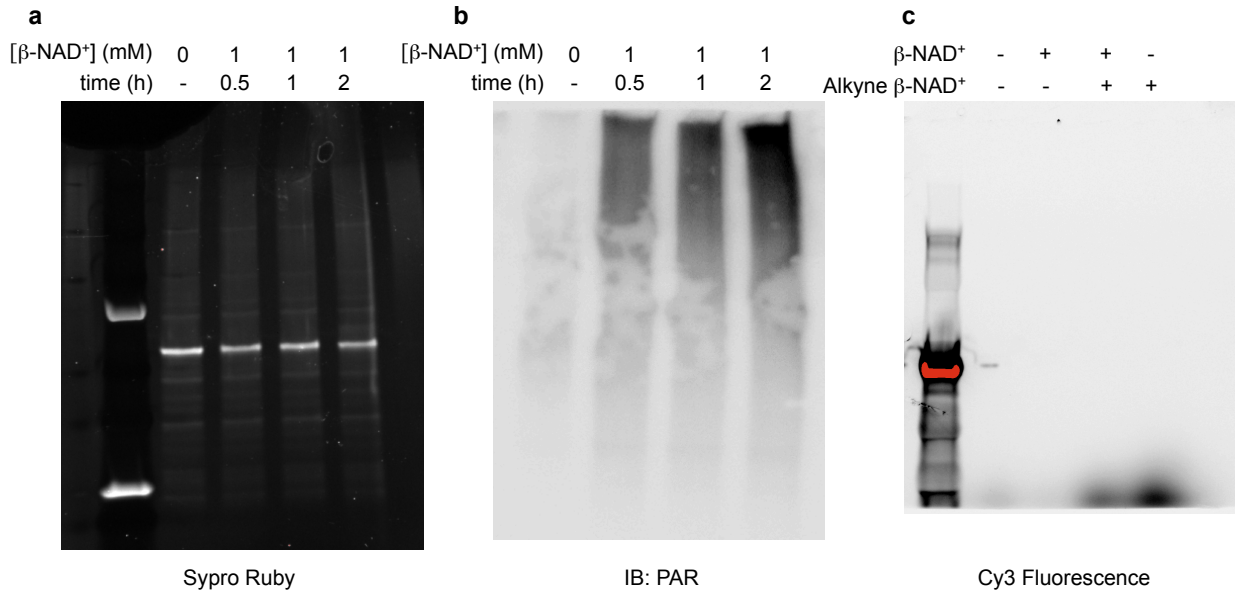


Figure 4.4. Assessment of PAR synthesis activity of the catalytic domain of tankyrase-1. a) Treatment of tankyrase-1 with natural β-NAD^+ does not cause noticeable retardation of protein migration by SDS-PAGE and sypro ruby staining. b) Some PAR formation can be seen by immunoblot. c) Incorporation of alkyne β-NAD^+ (**4-2**) is not seen by in gel fluorescence after Cy3 conjugation.

Having demonstrated low activity of tankyrase-1 and concluded that it is therefore unlikely to be ideal for chemoenzymatic PAR synthesis, we next assessed the ability of full-length human PARP-1 to incorporate β-NAD^+. Full-length human PARP-1 expressed in Sf21 insect cells was purchased from Enzo Life Sciences. Though purchase of the protein was not a sustainable solution due to its cost (10 μg/$292), preliminary assessment of its ability to incorporate natural and unnatural β-NAD^+ was desired. It was found that full length human PARP-1 robustly produced PAR as evident by retardation of migration of the protein by SDS-PAGE and coomassie staining (Figure 4.5a). PAR formation could also be seen by

immunoblotting for PARP-1 (Figure 4.5b). Finally, PARP-1 was able to incorporate alkyne β-NAD$^+$ (**4-2**) as evident by in gel fluorescence (Figure 4.5c). These experiments strongly suggested that full length PARP-1 would be an ideal enzyme for these studies, though its procurement presented a greater challenge.

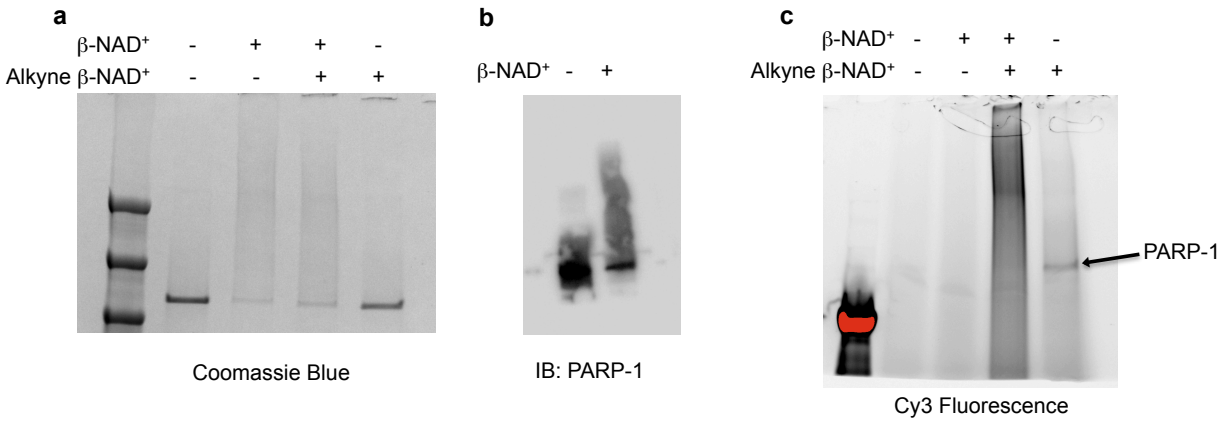


Figure 4.5. Assessment of activity of full-length human PARP-1. a) Treatment of PARP-1 with natural β-NAD$^+$ was well as a mixture with alkyne β-NAD$^+$ (**4-2**) results in decreased migration of PARP-1 by SDS-PAGE visualized by coomassie blue staining. b) The effect of a) can be seen by immunblot for PARP-1. c) Incorporation of alkyne β-NAD$^+$ (**4-2**) is seen by in gel fluorescence after Cy3 conjugation.

4.3 Expression, Purification and Activity of Full-length Human PARP-1

With full length human PARP-1 identified as the appropriate enzyme for the controlled synthesis of PAR polymers, the protein needed to be expressed so that it could be obtained in greater quantity and at lower cost than what was obtained from commercial sources. While often expressed in insect cells, limited reports of expression of PARP-1 in *E. coli* do exist[10]. Upon obtaining an appropriate plasmid from the Ahel Lab, His_6-PARP-1 was expressed and purified using a two step method involving Ni-NTA and subsequent Heparin chromatography[10] (Figure 4.6a,b). Upon concentration, PARP-1 expressed via this method was shown to be active both

towards polymerization of natural β-NAD$^+$ (Figure 4.6c), as well as towards incorporation of alkyne β-NAD$^+$ (**4-2**) (Figure 4.6d).

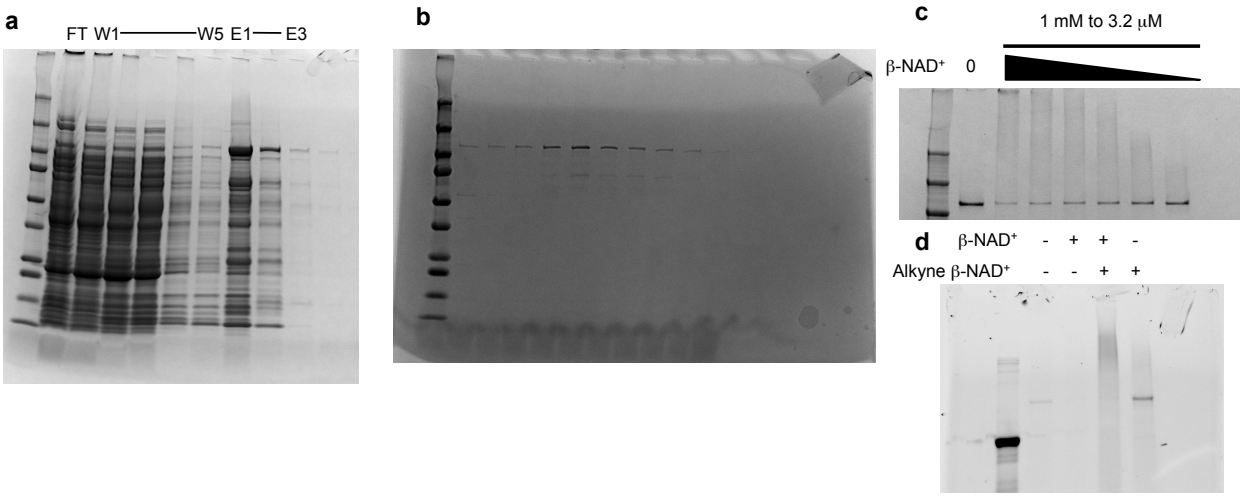


Figure 4.6. Expression and purification of full length human PARP-1 a) Initial purification by Ni-NTA chromatography. b) Fractions from heparin chromatography containing pure protein. c) Ability of expressed PARP-1 to incorporate natural β-NAD$^+$ as seen by coomassie staining. d) Expressed PARP-1 incorporates alkyne β-NAD$^+$ (**4-2**) as seen by Cy3 in gel fluorescence.

4.4. Attempts to Chemoenzymatically Synthesize PAR Oligomers Using Compound **4-1**

Having chemically synthesized β-NAD$^+$ analogs **4-1**, **4-2** and **4-3**, and having recombinantly expressed PARP-1 in acceptable yield, attempts to use these tools for the chemoenzymatic synthesis of short PAR oligomers were undertaken. A simple SDS-PAGE assay was developed to monitor β-NAD$^+$ incorporation (Figure 4.7). In this assay, PARP-1 was immobilized to magnetic streptavidin beads by reacting the protein with biotin maleimide via its free cysteine residues. The protein was then incubated with oNB β-NAD$^+$ (**4-1**) for 45 min. Next, excess substrate was removed by washing with buffer and the protein was treated with UV light to photochemically cleave the oNB protecting group on the ADP-ribosylated PARP-1. Subsequent treatment with β-NAD$^+$ alkyne (**4-2**) should in theory lead to incorporation at the

newly deprotected ADP-ribose sites and formation of PARP homogenously labeled with alkyne containing ADP-ribose dimer. Such a spices would then be labeled with Cy3 and visualized by in gel fluorescence.

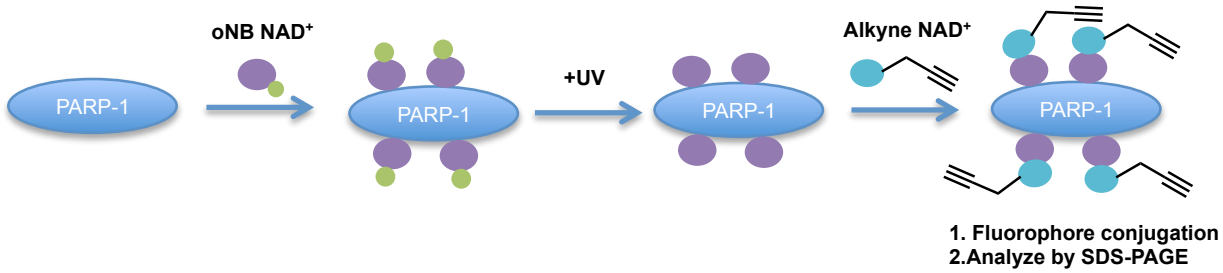


Figure 4.7. Workflow for SDS-PAGE assay involving unnatural β-NAD^+ analogs.

In practice, performance of the assay as described did not lead to robust formation of PARP labeled with the ADP-ribose dimer as expected. While a fluorescent band is seen when this process is carried out, it is also seen for the control where treatment with UV light is not carried out, suggesting that the fluorescent band seen is not solely due to the ADP-ribose dimer as expected (Figure 4.8a). While excess deoxy β-NAD^+ (**4-3**) is capable of inhibiting incorporation of alkyne β-NAD^+ (**4-2**) in a dose-dependent manner (Figure 4.8b), no conditions were found to elicit this effect for oNB β-NAD^+ (**4-1**). It was therefore concluded that oNB β-NAD^+ (**4-1**) was a poor substrate for PARP-1 and that its inability to block incorporation of alkyne β-NAD^+ (**4-2**) was the result of it not being incorporated itself.

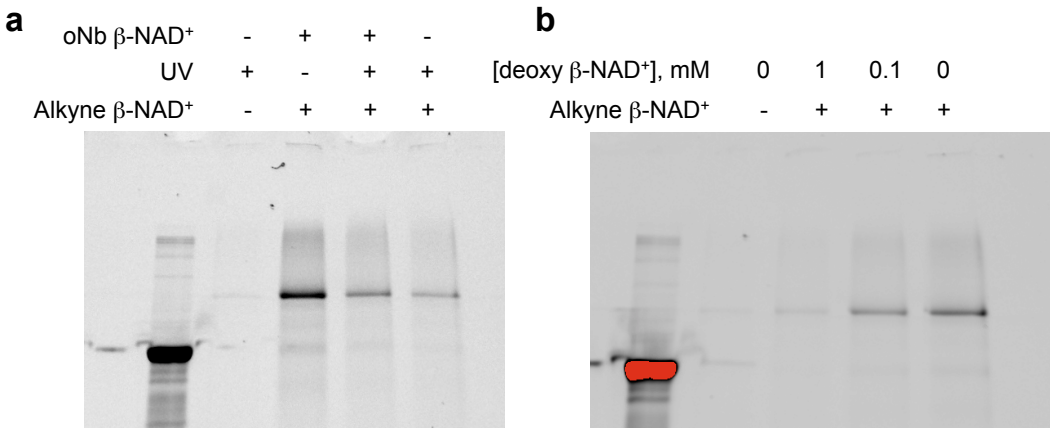


Figure 4.8. Chemoenzymatic synthesis of ADP-ribose oligomers by in gel fluorescence using oNB β-NAD$^+$ (**4-1**). a) oNB β-NAD$^+$ (**4-1**) does not block incorporation of alkyne β-NAD$^+$ (**4-2**). Alkyne β-NAD$^+$ (**4-2**) incorporation can be blocked by treatment with deoxy β-NAD$^+$.

4.5 Design of a Second Generation Photocleavable β-NAD$^+$ Analog

After determining that PARP-1 could not incorporate oNB β-NAD$^+$ (**4-1**) as a substrate, a new design was needed. That propargyl β-NAD$^+$ was incorporated by PARP-1 was encouraging, as it suggested that incorporation of analogs modified at 2' position of β-NAD$^+$ by PARP was possible. It was reasoned that because the propargyl group is substantially smaller than the oNB group, perhaps decreasing the size of the protecting group or moving it further away from the 2'-OH of β-NAD$^+$ could be a successful strategy. The [2(-nitrobenzyl)oxy] methyl group (nbm) was therefore chosen as a replacement for the orthonitrobenzyl group as used for compound **4-1**[15]. Upon photolysis of an analog containing this group (**4-9**), β-NAD$^+$ would be generated along with formaldehyde (Figure 4.9). We hypothesized that the methylene spacer of this compound would move steric bulk away from the enzyme, facilitating better incorporation of this analog.

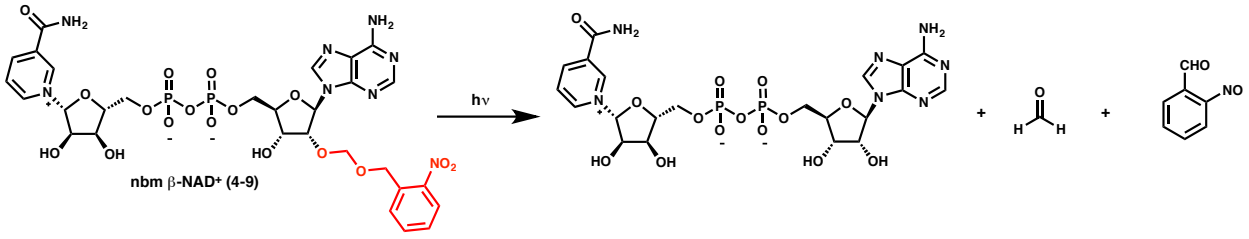


Figure 4.9. Second generation photochemical β-NAD$^+$ analog nbm β-NAD$^+$ (**4-9**) and its cleavage to β-NAD$^+$ and formaldehyde upon photolysis.

In the forward sense, nbm β-NAD$^+$ (**4-9**) was synthesized in a short sequence commencing from N-benzoyl-5'-dimethoxytrityl adenosine (Figure 4.10). Alkylation using a known procedure gave alcohol **4-11**[15]. Deprotection of the dimethoxytrityl group and benzoyl group was accomplished in a two step process[15]. First, treatment with formic acid removed the dimethoxytrityl group, and subsequent treatment with ammonia in methanol cleaved the N-benzoyl group to give compound **4-12**. Strategically, the installation of protecting groups to give **4-10** followed by their quick removal after alkylation to give **4-12** was less than desirable. Still, the protection of compound **4-11** did facilitate its purification relative to systems where the alkylation was performed without protecting groups and this strategy was therefore used. Phosphorylation of **4-12** proceeded smoothly to give **4-13** after C_{18} purification. Finally, phosphorimidazolide formation and coupling to NMN gave nbm β-NAD$^+$ (**4-9**) in only trace yields. The yields of **4-9** in this case are poor because of hydrolysis of the imidazolide under the reaction conditions forms **4-13**, which is challenging to separate from **4-9** even by preparative HPLC. However, the quantities afforded of **4-9** were sufficient to test in preliminary experiments.

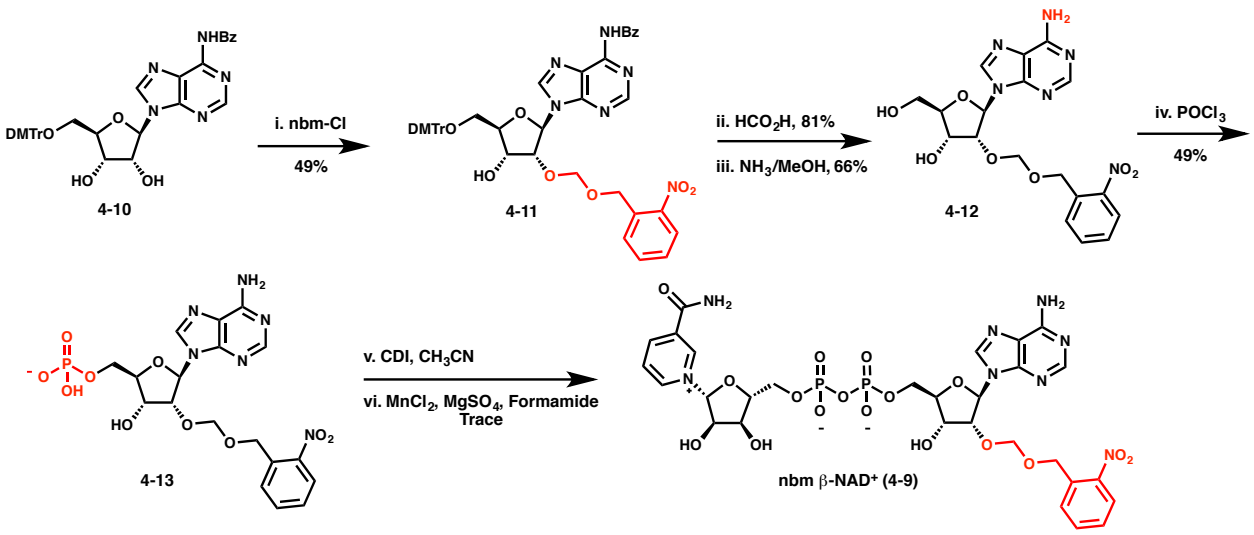


Figure 4.10. Synthesis of nbm β-NAD^+ (**4-9**). i. nbm-Cl (1.5 eq), Bu_2SnCl_2 (1.1 eq), EtN^iPr_2 (5 eq), $(CH_2Cl)_2$, 60°C, 2 h, 49%. ii. HCO_2H, $CHCl_3$, 0°C, 1 h, 81%. iii. NH_3/MeOH, rt to 40 °C, 24 h, 66%. iv. $POCl_3$, 0°C, 1 h, 49%. v. CDI, CH_3CN, rt, 60 h. vi. NMN, $MnCl_2$, $MgSO_4$, Formamide, rt, 18 h, trace for two steps.

4.6 Use of Second Generation nbm β-NAD^+ (**4-9**) for Chemoenzymatic Synthesis

Upon accessing the second generation photocleavable β-NAD^+ analog (**4-9**), its ability to perform in the chemoenzymatic synthesis assay described in section 4.4 was evaluated. Gratifyingly, incubation of this analog with PARP-1 was capable of blocking incorporation of alkyne β-NAD^+ (**4-2**) (Figure 4.11) minimally suggesting that the new analog is a better substrate for PARP-1 than the original oNB β-NAD^+ (**4-1**). Unfortunately, in contrast to what was seen in the oNB β-NAD^+ (**4-1**) incorporation experiments, PARP-1 was not able to append alkyne β-NAD^+ (**4-2**) upon photocleavage (Figure 4.11). While the exact reason for this inability to iterate is unknown, we suspected that solid phase attachment of PARP-1 combined with the aldehyde byproducts released upon photocleavage might be responsible.

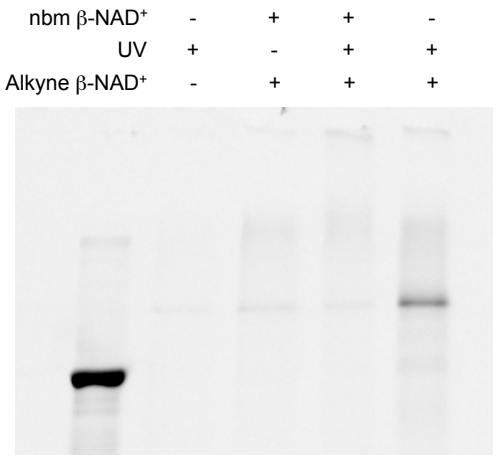


Figure 4.11. Chemoenzymatic synthesis of ADP-ribose oligomers by in-gel fluorescence using nbm β-NAD^+ (**4-9**).

4.7. Attempts to Perform Chemoenzymatic Synthesis via the *in trans* Activity of PARP-1

All previous approaches in this chapter have relied on the ability of PARP-1 to undergo automodification (*in cis* activity) and have used PARP-1 as the substrate on which PAR oligomers and polymers are built. Having not identified a good method by which to chemoenzymatically synthesize PAR oligomers via the automodification activity of PARP-1, we next sought to pursue an alternative chemoenzymatic synthesis strategy through PARP-1's ability to modify substrates *in trans*. A number of proteins, most commonly histones, are known to be *in trans* substrates for PARP-1[11,16]. Additionally, a few examples exist where short synthetic peptides derived from PARP-1 or other proteins have served as substrates for PAR synthesis mediated by PARP-1[17-19]. We reasoned that an advantage of such an approach, in contrast to the solid phase, would be the utilization of PARP-1 in solution. Use of PARP-1 in solution could potentially lead to increased enzymatic activity relative to when appended to the solid phase.

Synthetic peptides known to be substrates for PARP-1[17,19] were obtained and we incubated them with compounds **4-2** and/or **4-3**. Incorporation of these analogs was monitored

by LCMS. Unfortunately, these peptides were not labeled with the compounds under any conditions tested. While this result is somewhat surprising, literature reports suggesting that these peptides serve as acceptors for PARylation rarely report yields for such reactions and usually report automodification if it is detected by MS or immunoblot[17-19]. Thus, it is likely that the actual conversion and synthesis of PAR on these peptides is low in a preparative sense.

4.8. *In trans* Activity of PARP-1 Using the ADP-ribose Dimer (**2-1**) as a "Seed" for Polymerization

With the failure of the *in cis* (sections 4.4 and 4.6) and *in trans* (section 4.7) approaches to chemoenzymatic synthesis, we reevaluated our approach and more extensively considered the process of PAR synthesis by PARP-1. While to this point much attention had been focused on what protein or peptide substrates PARP-1 recognizes for PARylation, we reasoned that as PAR polymers become increasingly long, PARP-1 likely recognizes the polymer itself and not the substrate protein or peptide when adding units to the polymer. Therefore, we reasoned that free, non-protein bound PAR oligomers in solution could serve as substrates for further PARylation.

To test this hypothesis, the ADP-ribose dimer (**2-1**) was treated with alkyne β-NAD^+ (**4-2**) and PARP-1 and the mixture was analyzed by LCMS (Figure 4.12). Gratifyingly, small amounts of formation of the propargyl ADP-ribose trimer were seen, especially when Mg^{2+} ions were omitted from the reaction buffer (Figure 4.12b). Omission of Mg^{2+} has been reported to increase the *in trans* activity of PARP-1, while inclusion of Mg^{2+} is believed to bias PARP-1 towards automodification activity[20]. Attempts to optimize this reaction further by inclusion of putrescine in the reaction buffer were unsuccessful[20].

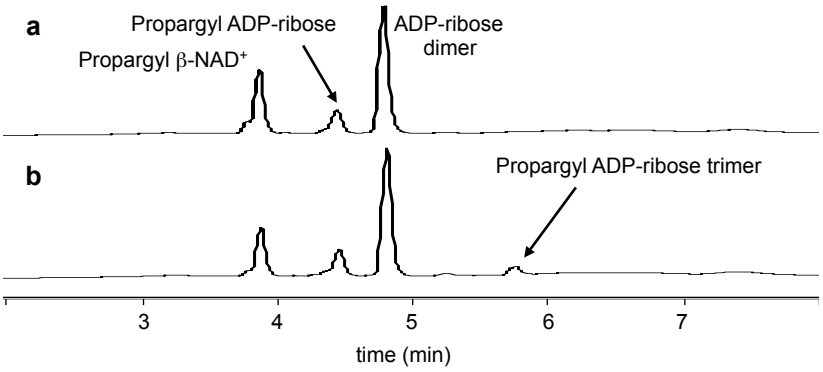


Figure 4.12. Formation of propargyl ADP-ribose trimer by treatment of ADP-ribose dimer (**2-1**) and propargyl β-NAD^+ with PARP-1. a) Standard buffer with Mg^{2+}. b) Buffer omitting Mg^{2+}. The peak at 5.88 in b) is consistent with the mass of the propargyl ADP-ribose by mass spectrometry.

With the exciting result in Figure 4.12b, attention was turned to investigating the generality of this approach, optimizing, and extending it to a system in which natural ADP-ribose oligomers and polymers could be synthesized. If natural β-NAD^+ could also perform this *in trans* activity and if such activity could be controlled, natural ADP-ribose oligomers could be produced. It was found that treatment of the ADP-ribose dimer (**2-1**) with β-NAD^+ could, in fact, effect formation of the ADP-ribose trimer (Figure 4.13). This activity was found to be concentration dependent, as high concentrations of β-NAD^+ (Figure 4.13a) failed to produce formation of the ADP-ribose trimer, while intermediate concentrations (Figure 4.13c) were more effective. The process does require PARP-1 (Figure 4.13c vs. Figure 4.13d). Additionally, the amount of the ADP-ribose trimer formed relative to the amount of propargyl ADP-ribose trimer formed in Figure 4.12 was greater, suggesting that β-NAD^+ is a better substrate than the unnatural propargyl β-NAD^+. It was also found that nbm β-NAD^+ (**4-9**) was inactive in this assay, as PARP-1 was unable to synthesize an nbm ADP-ribose trimer from the ADP-ribose dimer (**2-1**) and nbm β-NAD^+ (**4-9**) (data not shown). These results could explain the failure of

the *in cis* approach described previously, as the unnatural β-NAD^+ described above may simply be poor substrates for PARP-1 incorporation. Additionally, ADP-ribose dimer formation (**2-1**) could be detected from treatment of ADP-ribose with β-NAD^+ (data not shown), but the conversion was generally lower than what could be seen for ADP-ribose dimer (**2-1**) conversion to ADP-ribose trimer. Formation of the ADP-ribose tetramer in this assay was detected but was minimal. This result suggests that the low concentrations of trimer present in the reaction coupled with the low enzymatic efficiency of PARP-1 towards this process combine to not lead to runaway polymerization to longer oligomers.

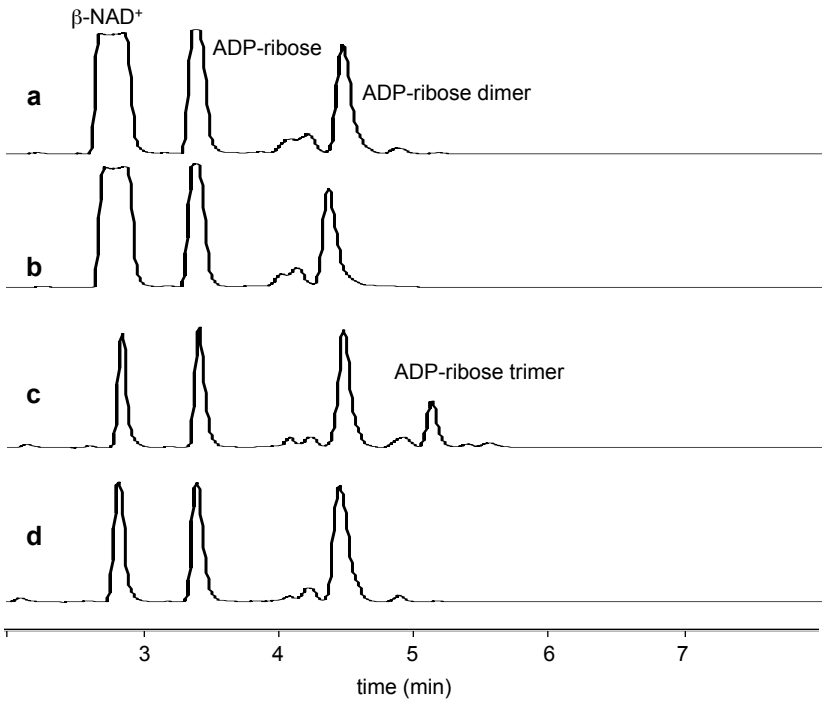


Figure 4.13. The ADP-ribose trimer can be formed from β-NAD^+ and the ADP-ribose dimer (**2-1**), though not at high concentrations of β-NAD^+ a) 5 mM concentration of β-NAD^+ with PARP-1 and the ADP-ribose dimer (**2-1**) b) 5 mM concentration of β-NAD^+ and the ADP-ribose dimer (**2-1**) without enzyme c) 500 μM concentration of β-NAD^+ with PARP-1 and the ADP-ribose dimer (**2-1**) d) 500 μM concentration of β-NAD^+ and the ADP-ribose dimer (**2-1**) without enzyme.

4.9. Preparative Synthesis of ADP-ribose Oligomers Using the "Seed" Method

The synthesis of the ADP-ribose trimer from the ADP-ribose dimer (**2-1**) as demonstrated in figure 4.13 was next scaled up to form meaningful quantities of the trimer and enable synthesis of even longer oligomers. Thirty fold scale-up of the conditions in figure 4.13 allowed for synthesis, purification, and isolation and characterization of the trimer by LCMS (Figure 4.14). It was found that the trimer could be isolated as a mostly pure substance, though non-insignificant amounts of a byproduct lacking the terminal ribose phosphate was formed. This product is not seen in the crude reaction (as in figure 4.13) and is likely a result of decomposition of the product upon concentration. Base lability of ADP-ribose oligomers as has been reported previously[4,21]. Modification of the conditions under which oligomers are isolated will likely resolve this decomposition issue and more cleanly provide trimer.

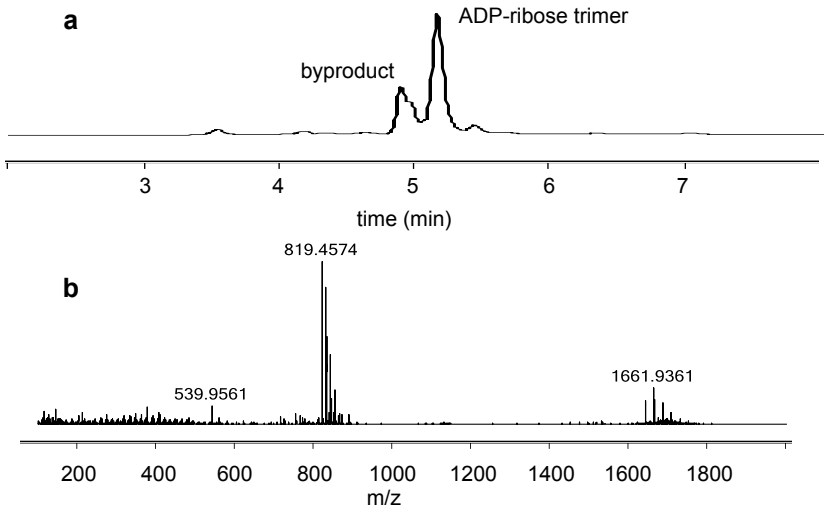


Figure 4.14. Isolation of the ADP-ribose trimer from the chemoenzymatic route. a) HPLC trace showing trimer but with some cleavage product. b) MS trace of major peak showing $[M-2H]^{2-}$ = 819, consistent with the ADP-ribose trimer.

4.10 Conclusions and Outlook

The chemoenzymatic synthesis of ADP-ribose oligomers as described in sections 4.8 and 4.9 represents a promising approach for the generation of PAR oligomers and polymers. Key to this approach is the ability of PARP-1 to modify substrates *in trans* with enough efficiency to make the reaction synthetically useful, but not so extensively that runaway polymerization is seen. As previous approaches to enzymatic synthesis of PAR oligomers and polymers have not started with an oligomer of defined length and used it as a substrate for synthesis, this ability of PARP-1 has not been previously recognized.

How PARP-1 is able to selectively carry out this process in a controlled manner without runaway polymerization is still not entirely clear. One explanation is simply the math of the process. For instance, assume that PARP-1 conversion of the ADP-ribose dimer to the corresponding trimer occurs in about 30% yield (this would be consistent with figure 4.13). Next, assume that further conversion of the trimer to the tetramer occurs in similar yield. Therefore, the yield of dimer to tetramer in this process would be: 30% x 30% = 9%. This number is smaller than but not grossly inconsistent with the amount of tetramer observed experimentally upon LCMS analysis. By extension of this analysis, the yield of the pentamer from the dimer would be about 3%, and therefore likely barely detectable.

In reality, the yield of the second step may be lower or higher than the yield of the first. The amount of trimer in solution available for automodification at 30% conversion (150 μM) is much reduced relative to both amount of dimer available at the beginning of the experiment (500 μM), and the amount of dimer available at 30% conversion to trimer (350 μM). Therefore, one could imagine that under the experimental conditions the propensity for conversion of trimer to tetramer is lower. However, preliminary data (not shown) suggests that conversion of the

concentrated trimer to the tetramer occurs in higher yield than conversion of the dimer to the trimer (and both occur better than conversion of ADP-ribose to the dimer). Experimentally, the first factor seems to play a larger role than the second, though how this process will progress for longer oligomers remains to be seen. Regardless, even the selective synthesis of narrow distributions of PAR would have advantages over enzymatic synthesis and separation of PAR of all lengths between 2 and 200.

Homogeneous synthesis of short PAR oligomers of defined length should allow debated questions about PAR biology to be answered. For example, the exo vs. endo preference of PARG is an unresolved issue in the literature[22,23]. Using the approach described in sections 4.8 and 4.9 we imagine being able to access the ADP-ribose tetramer to address this issue. If PARG is an enzyme with primarily endo activity, we would expect to see two molecules of the ADP-ribose dimer (**2-1**) upon PARG treatment of the tetramer. However, if PARG possesses primarily exo activity, we would expect to see the trimer and a molecule of ADP-ribose upon treatment. The use of short PAR oligomers (minimally the ADP-ribose tetramer) should resolve these questions.

4.11 Experimental

Expression and Purification of the catalytic domain of GST-Tankyrase-1

Expression and purification was performed essentially as described previously[14] with slight modifications. Ampicillin and chloramphenicol were used at concentrations of 100 μg/mL and 35 μg/mL, respectively. Expression was performed on a 1 L scale and IPTG was added at a concentration of 250 μM once the culture had reached OD_{600} of 0.85. Expression was carried out for 20 h at 15°C.

Expression and purification of full length human PARP-1

PARP-1 was expressed and purified essentially as described[10]. Briefly, the plasmid for full-length human PARP-1 was a gift from the Ahel Lab. The plasmid was transformed into *E. coli* Rosetta cells and grown overnight in 10 mL of LB containing 50 μg/mL and 35 μg/mL kanamycin and chloramphenicol, respectively. The overnight culture was added to 1 L of LB (also with 50 μg/mL kanamycin and 35 μg/mL chloramphenicol) and the culture was induced with 200 μM IPTG once it had reached an OD_{600} of 0.9. The expression was performed for 20 h at 16°C. Harvesting and lysis was performed as described. The protein was purified by Ni-NTA chromatography and Heparin chromatography in a 4°C room as described. The size exclusion step was omitted.

Tankyrase Activity Assay

To a buffered solution (100 mM Tris, 10 mM $MgCl_2$, 1 mM DTT pH = 7.6) of β-NAD^+ or β-NAD^+ (1 mM or as indicated) was added Tankyrase-1 to a final concentration of 5 μM. The reaction was incubated for the indicated time and SDS-loading dye was added and the sample heated to 95°C for 5 min. The sample was analyzed by SDS-PAGE on a 4-20% or 12% gel. The gel was stained with Sypro Ruby protein stain per the manufacturer's instructions.

For western blotting, the sample was transferred to a PVDF membrane and blocked in 5% milk for 1 h at room temperature. The membrane was incubated with anti-PAR primary antibody (1:1000 dilution in 5% milk), washed and incubated with the HRP-conjugated secondary antibody for 1 h. The blot was imaged using a ImageQuant LAS4010

For in gel fluorescence experiments, the above setup was employed, but with 100 μM NAD^+ or NAD^+ analog. After the indicated time, sulfo-Cy3-azide (1.4 mM), THPTA (200 mM),

$CuSO_4$ (4 mM) and sodium ascorbate (320 mM) were added and incubated for 1-2 h. The reactions were put through a illustra Microspin G-25 column and subjected to SDS-PAGE as described above. The gels were visualized using a ImageQuant LAS4010.

PARP-1 Activity Assay by SDS-PAGE

To a buffered solution (100 mM Tris, 10 mM $MgCl_2$, 1 mM DTT pH = 7.6) of β-NAD^+ or β-NAD^+ (1 mM or as indicated) was added calf thymus DNA to a final concentration of 100 μg/mL. Next, PARP-1 to a final concentration of 150 nM. The reaction was incubated for the indicated time and SDS-loading dye was added and the sample heated to 95^{o}C for 5 min. The sample was analyzed by SDS-PAGE on a 4-20% or 12% gel and stained with Imperial protein stain per the manufacturer's instructions.

For western blotting, the sample was transferred to a PVDF membrane and blocked in 5% milk for 1 h at room temperature. The membrane was incubated with anti-PARP-1 primary antibody (1:1000 dilution in 5% milk), washed and incubated with the HRP-conjugated secondary antibody for 1 h. The blot was imaged using a ImageQuant LAS4010.

For in gel fluorescence experiments, the above experimental conditions were employed and the procedure for visualization was the same as for Tankyrase-1.

PARP-1 immobilization and in gel activity assay

PARP-1 (10 μM) was treated with biotin maleimide (100 μM) in PARP activity buffer. After 30 min, streptavidin magnetic beads (50 μL) were added and the reaction components (as above) were added. The reaction was incubated for 40 min at which point the beads were washed 3x with PARP-1 buffer. The beads were resupsended in PARP-1 buffer and photocleaved for 40

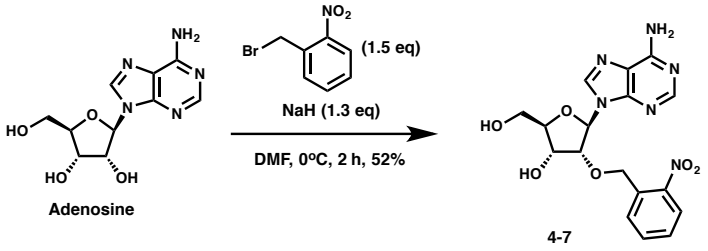
min under UV light. The beads were again washed and incubated with the new reaction components for 30 min. The beads were washed and incubated with the click chemistry components as above for 1 h. The beads were washed, SDS-loading dye was added, the samples were heated at 95°C for 5 min and the samples were analyzed by SDS-PAGE as described above.

Synthesis of ADP-Ribose Trimer from ADP-Ribose Dimer (2-1)

To a solution containing β-NAD^+ (500 μM or as indicated) and the ADP-ribose dimer (**2-1**) (500 μM) was added PARP buffer (20 mM Tris, 50 mM NaCl, 100 μM TCEP) and calf thymus DNA (100 μg/mL). PARP-1 (1 μM) was added and the reaction was heated at 37°C for 48 h. The final volume of the reaction was 10 μL. For analysis, the reaction was diluted to a final volume of 30 μL and analyzed by LCMS.

For preparative reactions, this process was scaled up to a final volume of 100 μL (component concentrations as above) and three reactions were run in parallel. The reactions were purified using HPLC, and fractions containing product were concentrated.

Chemical Synthesis



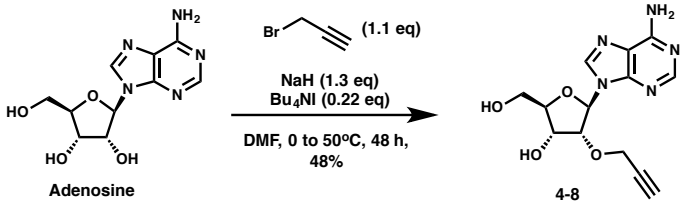
Compound 4-7: The general procedure of Ohtuska, Tanaka and Ikehara was followed[7]. To a round bottom flask was added adenosine (1.0 g, 3.74 mmol, 1 eq) and DMF (35 mL). The solution was heated to 55°C for 15 min at which point it was cooled to 0°C and sodium hydride

(186 mg, 4.86 mmol, 1.3 eq) was added. The solution was stirred at 0°C and 2-nitrobenzyl bromide (1.21 g, 5.61 mmol, 1.5 eq) was added and the reaction was stirred at 0°C for 2 h. The reaction was poured into ice water (400 mL), filtered and washed two times with ethanol to provide compound **4-7** (781 mg, 52%).

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.35 (s, 1H, H-2), 8.09 (s, 1H, H-8), 7.98 (dd, *J* = 8.1, 1.2 Hz, 1H), 7.65 – 7.55 (m, 2H), 7.52 – 7.45 (m, 1H), 7.35 (bs, 2H, -NH_2), 6.10 (d, *J* = 6.1 Hz, 1H, H-1'), 5.49 – 5.36 (m, 2H, -OH), 5.03 (d, *J* = 14.7 Hz, 1H), 4.86 (d, *J* = 14.7 Hz, 1H), 4.59 (dd, *J* = 6.1, 4.7 Hz, 1H, H-2'), 4.41 (td, *J* = 5.0, 3.1 Hz, 1H, H-3'), 4.04 (q, *J* = 3.4 Hz, 1H, H-4'), 3.69 (dt, *J* = 12.2, 4.2 Hz, 1H, H-5'), 3.58 (ddd, *J* = 12.2, 7.1, 3.6 Hz, 1H, H-5').

^{13}C NMR (126 MHz, DMSO-d_6) δ 156.2, 152.4, 149.0, 147.0, 139.6, 134.0, 133.6, 128.5, 124.4, 119.4, 86.4, 86.0, 81.1, 68.9, 67.8, 61.4.

HRMS (ESI) Calculated for $C_{17}H_{19}N_6O_6$ [$(M+H)^+$] 403.1366, found 403.1365



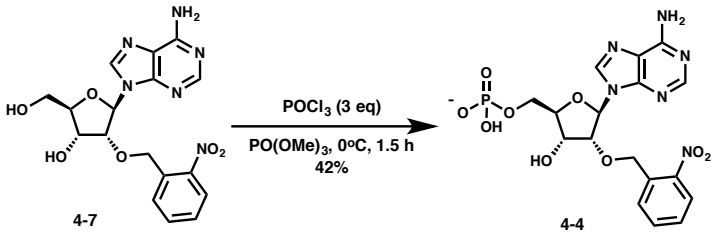
Compound 4-8: The general procedure of Madder *et al.* was followed and the product was consistent with that previously reported[8]. To a 0°C solution of adenosine (1.0 g, 3.74 mmol, 1 eq) and tetrabutylammonium iodide (303 mg, 0.82 mmol, 0.22 eq) was added sodium hydride (190 mg, 4.97 mmol, 1.3 eq). Propargyl bromide (498 mg, 4.19 mmol, 1.1 eq, 80% solution in toluene) was added and the reaction was heated to 50°C for 48 h. The compound was purified by silica column chromatography to afford compound 4-8 (544 mg, 48%).

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.37 (s, 1H, H-2), 8.15 (s, 1H, H-8), 7.36 (s, 2H, -NH_2), 6.03 (d, *J* = 6.3 Hz, 1H, H-1'), 5.49 (dd, *J* = 7.1, 4.6 Hz, 1H, 5'-OH), 5.35 (d, *J* = 5.1 Hz, 1H, 3',-

OH), 4.70 (dd, *J* = 6.4, 4.8 Hz, 1H, H-2'), 4.44 – 4.31 (m, 2H, H-3'), 4.28 (dd, *J* = 15.9, 2.3 Hz, 1H, C*H*$_2$-CCH), 4.20 (dd, *J* = 15.9, 2.3 Hz, 1H, C*H*$_2$-CCH), 4.01 (appq, *J* = 2.9 Hz, 2H, H-4'), 3.69 (appdt, *J* = 12.1, 4.1 Hz, 1H, H-5'), 3.57 (ddd, *J* = 12.3, 7.1, 3.5 Hz, 1H, H-5'), 3.34 (t, J = 2.5 Hz, 1H, CH$_2$-CC*H*)

^{13}C NMR (126 MHz, DMSO-d_6) δ 156.2, 152.5, 149.0, 139.8, 119.3, 86.7, 85.9, 79.7, 79.6, 77.6, 68.9, 61.6, 56.9.

HRMS (ESI) Calculated for $C_{13}H_{16}N_5O_4$ [$(M+H)^+$] 306.1202, found 306.1207



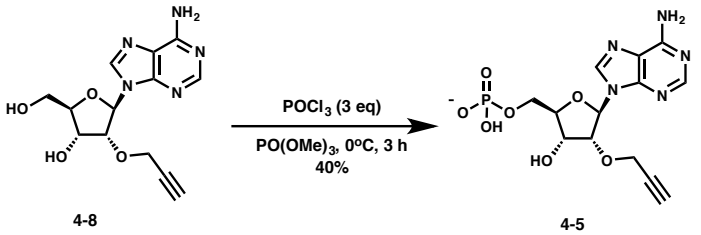
Compound 4-9: To a 0^{o}C solution of compound **4-7** (100 mg, 0.25 mmol, 1 eq) in trimethyl phosphate (0.5 mL) was added $POCl_3$ (0.070 mL, 0.746 mmol, 3 eq) and the reaction was stirred at 0^{o}C for 1.5 h. The reaction was quenched with water at 4^{o}C and purified by reverse phase chromatography to give compound **4-9** (50 mg, 42%).

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.52 (s, 1H, H-2), 8.16 (s, 1H, H-2), 7.89 (dd, *J* = 8.1, 1.3 Hz, 1H), 7.67 (dd, *J* = 7.9, 1.5 Hz, 1H), 7.50 (td, *J* = 7.6, 1.4 Hz, 1H), 7.39 (ddd, *J* = 8.1, 7.4, 1.5 Hz, 1H), 6.24 (d, *J* = 5.9 Hz, 1H, H-1'), 5.14 (d, *J* = 14.5 Hz, 1H), 4.98-4.95 (m, 1H), 4.69 (dd, *J* = 5.9, 4.9 Hz, 1H, H-2'), 4.61 (dd, *J* = 4.9, 3.3 Hz, 1H, H-3'), 4.31-4.27 (m, 1H, H-4'), 4.12 (m, 2H, H-5'), 3.07 (q, *J* = 7.3 Hz, 12H), 1.25 (t, *J* = 7.3 Hz, 18H).

^{13}C NMR (126 MHz, CD_3OD) δ 157.2, 153.8, 150.7, 148.8, 141.1, 135.2, 134.5, 130.1, 129.4, 125.4, 120.1, 87.3 (d, *J* = 8.9 Hz), 86.6, 84.1, 71.2, 70.1, 65.5 (d, *J* = 5.5 Hz), 47.1, 9.3.

^{31}P NMR (202 MHz, CD_3OD) δ 1.91

HRMS (ESI) Calculated for $C_{17}H_{20}N_6O_9P$ [$(M+H)^+$] 483.1029, found 483.1031



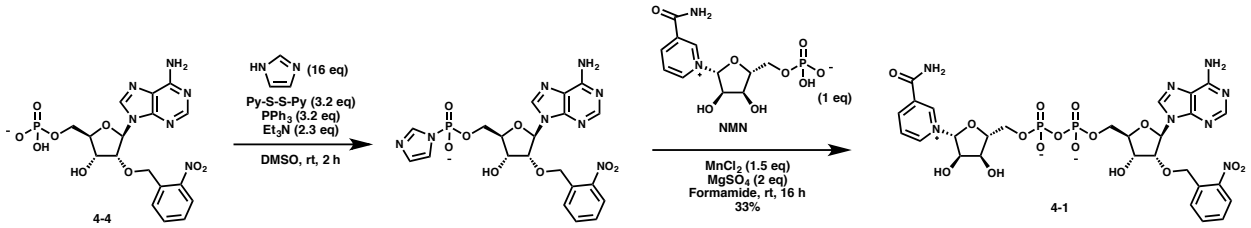
4-8 4-5

Compound 4-5: To a 0°C solution of compound **4-8** (50 mg, 0.164 mmol, 1 eq) in trimethylphosphate (0.33 mL) was added $POCl_3$ (0.045 mL, 0.491 mmol, 3 eq) and the reaction was stirred at 0°C for 3 h. The reaction was quenched with cold water and purified by reverse phase chromatography to afford compound **4-5** (27 mg, 40%).

^{1}H NMR (500 MHz, D_2O) δ 8.59 (s, 1H, H-2), 8.39 (s, 1H, H-8), 6.22 (d, *J* = 5.5 Hz, 1H, H-1'), 4.72 (dd, *J* = 5.5, 5.1 Hz, 1H, H-2'), 4.63 (dd, *J* = 5.1, 3.7 Hz, 1H, H-3'), 4.36 (ddd, *J* = 3.7, 2.4, 2.4 Hz, 1H, H-4'), 4.32 (dd, *J* = 16.1, 2.4 Hz, 1H, -C*H*$_2$CCH), 4.28 (dd, *J* = 16.1, 2.4 Hz, 1H, -C*H*$_2$CCH), 4.16 (ddd, *J* = 11.7, 4.9, 2.8 Hz, 1H, H-5'a), 4.11 (ddd, *J* = 11.7, 5.4, 3.2 Hz, 1H, H-5'b). 2.68 (t, *J* = 2.4 Hz, 1H, -CH$_2$CC*H*).

^{13}C NMR (126 MHz, D_2O) δ 149.8, 148.3, 144.7, 142.5, 118.4, 86.4, 84.8 (d, *J* = 8.6 Hz), 81.5, 78.7, 76.3, 69.2, 64.41 (d, *J* = 5.1 Hz), 58.3.

^{31}P NMR (202 MHz, CD_3OD) δ 0.12

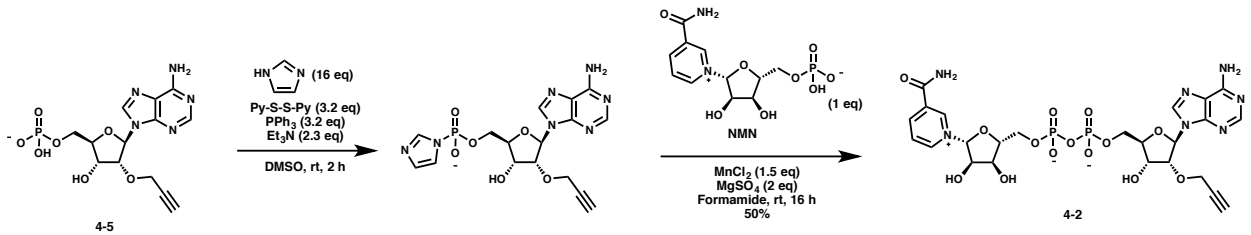


4-4 4-1

Compound 4-1: To a solution of compound **4-4** (50 mg, 0.104 mmol, 1 eq) in DMSO (1.2 mL) was added 2,2'-Dithiopyridine (73.1 mg, 0.332 mmol, 3.2 eq), triphenylphosphine (87.1 mg, 0.332 mmol, 3.2 eq), imidazole (113 mg, 1.664 mmol, 16 eq), and triethylamine (0.033 mL, 0.239 mmol, 2.3 eq). The reaction was stirred at room temperature for 2 h and purified by

reverse phase chromatography. To the isolated imidazolide (10.0 mg, 0.018 mmol, 1.2 eq) was added NMN (5.0 mg, 0.015 mmol, 1 eq) and formamide (0.2 mL). To the solution was added $MnCl_2$ (2.8 mg, 0.0225 mmol, 1.5 eq) and $MgSO_4$ (3.6 mg, 0.030 mmol, 2 eq). The reaction was stirred at room temperature for 16 h and purified by ion-pairing chromatography to give compound 4-1 (4.0 mg, 33%).

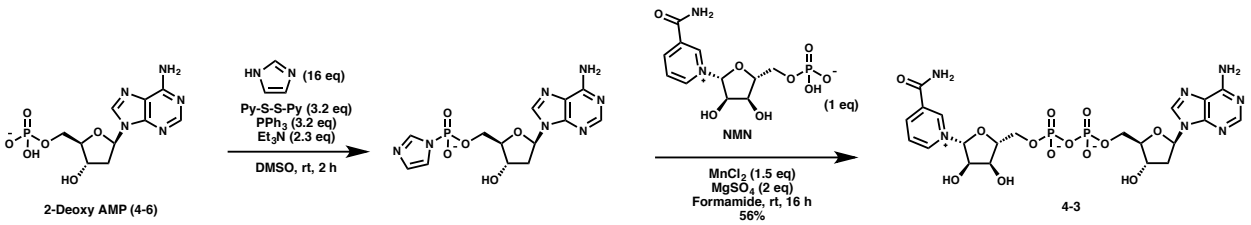
^{1}H NMR (500 MHz, D_2O) δ 9.14 (s, 1H), 8.95 (d, *J* = 6.2 Hz, 1H), 8.68 – 8.60 (m, 1H), 8.05 – 7.92 (m, 3H), 7.86 (s, 1H), 7.45 (d, *J* = 8.4 Hz, 1H), 7.29 – 7.21 (m, 1H), 7.13 – 7.06 (m, 1H), 5.89 (d, *J* = 5.4 Hz, 1H), 5.69 (d, *J* = 7.5 Hz, 1H), 5.12 (d, *J* = 11.7 Hz, 1H), 5.07 (d, *J* = 12.1 Hz, 1H), 4.59 – 4.42 (m, 3H), 4.37 (s, 1H), 4.33 – 4.17 (m, 3H), 4.05 (tt, *J* = 15.3, 4.7 Hz, 3H).



Compound 4-2: To a solution of compound **4-5** (27 mg, 0.066 mmol, 1 eq) in DMSO (0.8 mL) was added 2,2'-Dithiopyridine (47 mg, 0.21 mmol, 3.2 eq), triphenylphosphine (56 mg, 0.21 mmol, 3.2 eq), imidazole (72 mg, 1.06 mmol, 16 eq), and triethylamine (0.021 mL, 0.15 mmol, 2.3 eq). The reaction was stirred at room temperature for 2 h and purified by reverse phase chromatography. To the isolated imidazolide (5.5 mg, 0.012 mmol, 1.2 eq) was added NMN (3.4 mg, 0.010 mmol, 1 eq) and formamide (0.2 mL). To the solution was added $MnCl_2$ (1.9 mg, 0.015 mmol, 1.5 eq) and $MgSO_4$ (2.4 mg, 0.020 mmol, 2 eq). The reaction was stirred at room temperature for 16 h and purified by ion-pairing chromatography to give compound **4-2** (3.5 mg, 50%).

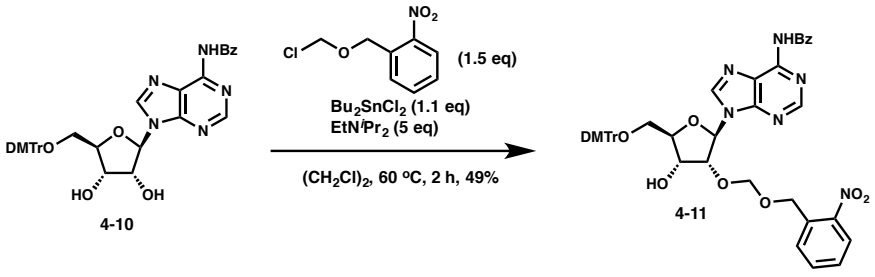
^{1}H NMR (500 MHz, D_2O) δ 9.17 (s, 1H), 8.98 (d, *J* = 6.2 Hz, 1H), 8.67 (d, *J* = 8.1 Hz, 1H), 8.31 (s, 1H), 8.06 – 7.96 (m, 2H), 5.96 (d, *J* = 6.3 Hz, 1H), 5.91 (d, *J* = 5.3 Hz, 1H), 4.64 (d, *J* =

6.2 Hz, 2H), 4.51 (dd, *J* = 5.3, 3.1 Hz, 1H), 4.41 – 4.35 (m, 1H), 4.32 (t, *J* = 5.2 Hz, 1H), 4.28 (dt, *J* = 5.4, 3.2 Hz, 1H), 4.25 – 4.19 (m, 2H), 4.15 (dd, *J* = 9.4, 2.3 Hz, 2H), 4.13 – 4.00 (m, 4H), 2.46 (t, *J* = 2.3 Hz, 1H).



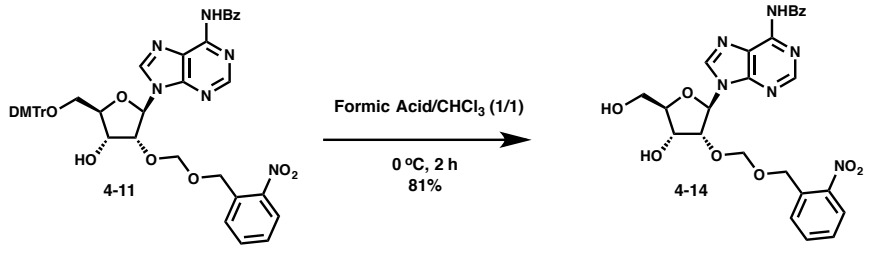
Compound 4-3: To a solution of 2'-deoxyadenosine 5'-monophosphate (**4-6**) (100 mg, 0.30 mmol, 1 eq) in DMSO (3.5 mL) was added 2,2'-Dithiopyridine (214 mg, 0.97 mmol, 3.2 eq), triphenylphosphine (254 mg, 0.97 mmol, 3.2 eq), imidazole (328 mg, 4.8 mmol, 16 eq), and triethylamine (0.097 mL, 0.70 mmol, 2.3 eq). The reaction was stirred at room temperature for 2 h and purified by precipitation from NaI in acetone. To the isolated imidazolide (7.3 mg, 0.018 mmol, 1.2 eq) was added NMN (5.0 mg, 0.015 mmol, 1 eq) and formamide (0.2 mL). To the solution was added $MnCl_2$ (2.8 mg, 0.0225 mmol, 1.5 eq) and $MgSO_4$ (3.6 mg, 0.030 mmol, 2 eq). The reaction was stirred at room temperature for 16 h and purified by ion-pairing chromatography.

^{1}H NMR (500 MHz, D_2O) δ 9.30 (s, 1H), 9.10 (d, *J* = 6.3 Hz, 1H), 8.78 (d, *J* = 8.2 Hz, 1H), 8.38 (s, 1H), 8.14 (dd, *J* = 8.1, 6.3 Hz, 1H), 8.11 (s, 1H), 6.40 (appt, *J* = 6.9 Hz, 1H), 6.05 (d, *J* = 5.5 Hz, 1H), 4.71 (dt, *J* = 6.3, 3.2 Hz, 1H), 4.51 (t, *J* = 2.6 Hz, 1H), 4.46 (t, *J* = 5.3 Hz, 1H), 4.40 (dd, *J* = 5.1, 2.7 Hz, 1H), 4.32 (ddd, *J* = 12.0, 4.2, 2.3 Hz, 1H), 4.28 – 4.24 (m, 1H), 4.23-4.09 (m, 3H), 2.80 (ddd, *J* = 13.8, 7.5, 6.0 Hz, 1H), 2.54 (ddd, *J* = 14.1, 6.4, 3.4 Hz, 1H).



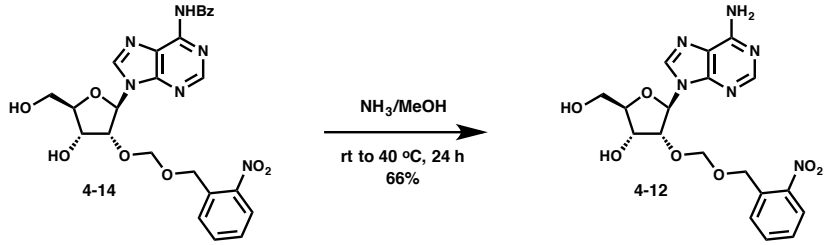
Compound 4-11: To a solution of compound **4-10** (500 mg, 0.742 mmol, 1 eq) in 1,2-dichloroethane (7 mL) was added EtN^iPr_2 (0.65 mL, 3.71 mmol, 5 eq) and Bu_2SnCl_2 (248 mg, 0.816 mmol, 1.1 eq). The reaction was stirred at room temperature for 1 h. After 1 h, 1-((chloromethoxy)methyl)-2-nitrobenzene (224 mg, 1.1 mmol, 1.5 eq) was added and the reaction was stirred at 60 ^oC for 1 h. The solution was quenched with saturated aqueous ammonium chloride, extracted three times with chloroform, dried through sodium sulfate and concentrated. The crude reaction was purified by silica column chromatography to afford compound **4-11** (305 mg, 49%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.94 (s, 1H), 8.62 (s, 1H), 8.16 (s, 1H), 8.02 (ddd, *J* = 8.7, 5.2, 1.5 Hz, 3H), 7.66 – 7.58 (m, 1H), 7.56 – 7.50 (m, 3H), 7.49 – 7.37 (m, 4H), 7.34 – 7.30 (m, 4H), 7.25 – 7.15 (m, 1H), 6.86 – 6.75 (m, 4H), 6.24 (d, *J* = 5.2 Hz, 1H), 5.12 (t, *J* = 5.1 Hz, 1H), 4.99 (s, 2H), 4.87 (d, *J* = 14.9 Hz, 1H), 4.81 (d, *J* = 14.8 Hz, 1H), 4.55 (q, *J* = 4.4 Hz, 1H), 4.29 (q, *J* = 3.9 Hz, 1H), 3.77 (s, 6H), 3.53 (dd, *J* = 10.6, 3.5 Hz, 1H), 3.42 (dd, *J* = 10.6, 4.1 Hz, 1H), 2.82 (d, *J* = 4.5 Hz, 1H).



Compound 4-14: To a 0 °C solution of compound **4-11** (280 mg, 0.33 mmol, 1 eq) in chloroform (3.3 mL) was added formic acid (3.3 mL). The reaction was stirred at 0 °C for 2 h. The compound was purified by silica column chromatography to afford compound **4-14** (145 mg, 81%).

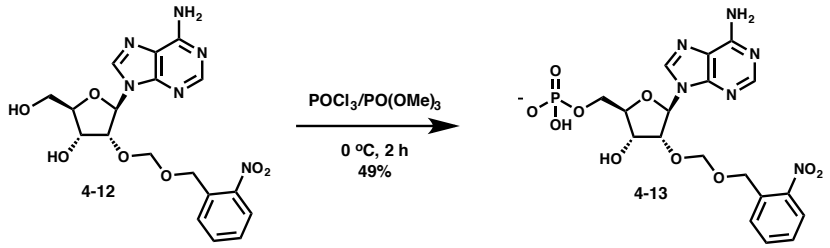
^{1}H NMR (500 MHz, CD_3OD) δ 8.63 (s, 1H), 8.51 (s, 1H), 8.13 – 8.05 (m, 2H), 7.95 (dd, *J* = 8.2, 1.3 Hz, 1H), 7.68 (t, *J* = 7.4 Hz, 1H), 7.59 (t, *J* = 7.7 Hz, 2H), 7.50 (td, *J* = 7.6, 1.4 Hz, 1H), 7.40 – 7.28 (m, 2H), 6.26 (d, *J* = 6.5 Hz, 1H), 5.01 (dd, *J* = 6.5, 4.9 Hz, 1H), 4.97 (d, *J* = 7.1 Hz, 1H), 4.91 (d, *J* = 7.0 Hz, 1H), 4.67 (d, *J* = 15.1 Hz, 1H), 4.62 (d, *J* = 15.0 Hz, 1H), 4.50 (dd, *J* = 5.0, 2.7 Hz, 1H), 4.21 (q, *J* = 2.8 Hz, 1H), 3.90 (dd, *J* = 12.4, 3.0 Hz, 1H), 3.80 (dd, *J* = 12.4, 3.0 Hz, 1H).



Compound **4-12**: To a solution of compound **4-14** (131 mg, 0.244 mmol, 1 eq) in methanol (3 mL) was added ammonia in methanol (3.5 mL of a 7M solution). The reaction was heated to 40 °C and stirred for 24 h. The solution was concentrated and purified by silica column chromatography to afford compound **4-12** (70 mg, 66%).

^{1}H NMR (500 MHz, *d_6*-DMSO) δ 8.30 (s, 1H), 7.98 (s, 1H), 7.96 (dd, *J* = 8.2, 1.3 Hz, 1H), 7.56 (td, *J* = 7.5, 1.3 Hz, 1H), 7.48 – 7.43 (m, 1H), 7.30 (dd, *J* = 7.8, 1.6 Hz, 1H), 7.24 (s, 1H), 6.02 (d, *J* = 6.5 Hz, 1H), 4.84 – 4.78 (m, 2H), 4.74 (d, *J* = 6.9 Hz, 1H), 4.71 (d, *J* = 14.6 Hz, 1H), 4.52 (d, *J* = 14.5 Hz, 1H), 4.29 (dd, *J* = 4.9, 2.8 Hz, 1H), 3.97 (appq, *J* = 3.3 Hz, 1H), 3.64 (dd, *J* = 12.3, 3.7 Hz, 1H), 3.52 (dd, *J* = 12.2, 3.6 Hz, 1H).

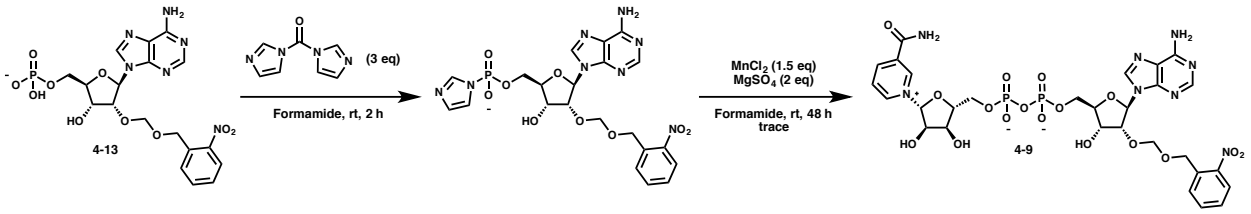
^{13}C NMR (126 MHz, d_6-DMSO) δ 155.9, 152.3, 148.8, 146.7, 139.7, 133.7, 133.4, 128.4, 128.2, 124.4, 119.2, 94.1, 86.4, 86.2, 78.4, 69.3, 65.6, 61.4.



Compound 4-13: To a solution of compound **4-12** (20 mg, 0.046 mmol, 1 eq) in trimethyl phosphate (0.3 mL) was added phosphorous oxychloride (0.013 mL, 0.139 mmol, 3 eq). The reaction was stirred at 0 ^{o}C for 2 h and purified by reverse phase chromatography to afford compound **4-13** (14 mg, 49%).

^{1}H NMR (500 MHz, CD_3OD) δ 8.57 (s, 1H), 7.99 – 7.90 (m, 2H), 7.53 (td, *J* = 7.6, 1.3 Hz, 1H), 7.44 – 7.37 (m, 1H), 7.29 (dt, *J* = 7.9, 1.1 Hz, 1H), 6.21 (d, *J* = 6.8 Hz, 1H), 5.01 (d, *J* = 7.0 Hz, 1H), 4.92-4-88 (m, 2H), 4.68 (d, *J* = 15.2 Hz, 1H), 4.61 (d, *J* = 15.2 Hz, 1H), 4.54 (dd, *J* = 5.0, 2.1 Hz, 1H), 4.30 – 4.21 (m, 1H), 4.15 – 4.07 (m, 2H).

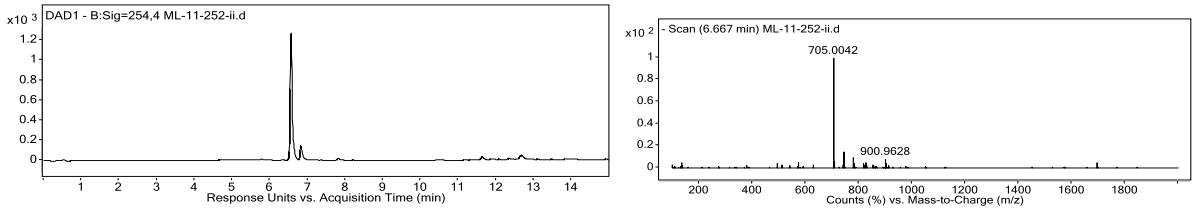
^{31}P NMR (202 MHz, CD_3OD) δ 1.98



Compound 4-9: To a solution of compound **4-13** (7.0 mg, 0.011 mmol, 1 eq) in formamide (0.2 mL) was added carbonyldiimidazole (5.5 mg, 0.034 mmol, 3 eq). The reaction was stirred at room temperature for 2 h and quenched with 0.5 mL of 1% Et_3N in MeOH (v/v). The solution was concentrated and evaporated three times with dry acetonitrile. The solution was dissolved in formamide (0.2 mL) and NMN (3.8 mg, 0.011 mmol, 1 eq) was added. $MnCl_2$ (2.2 mg, 0.017

mmol, 1.5 eq) and $MgSO_4$ (2.7 mg, 0.023 mmol, 2 eq) were added and the reaction was stirred at room temperature for 48 h. The product was purified by preparative HPLC. A small amount was isolated and a 1 mM solution was formed based on absorbance compared to compound **4-1**.

LCMS:



β-NAD^+ analogs fragment with loss of nicotinamide by ESI m/z = 705.

4.12 References

1. Kistemaker, H. A. V. *et al.* Synthesis of well-defined adenosine diphosphate ribose oligomers. *Angew. Chem. Int. Ed. Engl.* **54,** 4915–4918 (2015).
2. Kistemaker, H. A. V., Meeuwenoord, N. J., Overkleeft, H. S., van der Marel, G. A. & Filippov, D. V. On the Synthesis of Oligonucleotides Interconnected through Pyrophosphate Linkages. *Eur. J. Org. Chem* **2015,** 6084–6091 (2015).
3. Wang, Y., Rösner, D., Grzywa, M. & Marx, A. Chain-Terminating and Clickable NAD^+ Analogues for Labeling the Target Proteins of ADP-Ribosyltransferases. *Angew. Chem. Int. Ed.* **53,** 8159–8162 (2014).
4. Tan, E. S., Krukenberg, K. A. & Mitchison, T. J. Large-scale preparation and characterization of poly(ADP-ribose) and defined length polymers. *Anal. Biochem.* **428,** 126–136 (2012).
5. Wang, P. Photolabile Protecting Groups: Structure and Reactivity. *Asian J. Org. Chem* **2,** 452–464 (2013).
6. Lee, J. *et al.* A chemical synthesis of nicotinamide adenine dinucleotide (NAD+). *Chem. Commun.* **0,** 729–730 (1999).
7. Ohtsuka, E., Tanaka, S. & Ikehara, M. Studies on transfer ribonucleic acids and related compounds. XVI. Synthesis of ribooligonucleotides using a photosensitive o-nitrobenzyl protection for the 2'-hydroxyl group. *Chem. Pharm. Bull* **25,** 949–959 (1977).
8. Jawalekar, A. M., de Beeck, M. O., van Delft, F. L. & Madder, A. Synthesis and incorporation of a furan -modified adenosine building block for DNA interstrand crosslinking. *Chem. Commun.* **47,** 2796–2798 (2011).
9. Kraus, W. L. & Hottiger, M. O. PARP-1 and gene regulation: Progress and puzzles. *Molecular Aspects of Medicine* **34,** 1109–1123 (2013).
10. Langelier, M.-F., Planck, J. L., Servent, K. M. & Pascal, J. M. in *Poly(ADP-ribose) Polymerase* **780,** 209–226 (Humana Press, 2011).
11. Altmeyer, M., Messner, S., Hassa, P. O., Fey, M. & Hottiger, M. O. Molecular mechanism

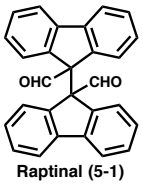
of poly(ADP-ribosyl)ation by PARP1 and identification of lysine residues as ADP-ribose acceptor sites. *Nucleic Acids Res.* **37,** 3723–3738 (2009).
12. Langelier, M.-F., Planck, J. L., Roy, S. & Pascal, J. M. Structural Basis for DNA Damage–Dependent Poly(ADP-ribosyl)ation by Human PARP-1. *Science* **336,** 728–732 (2012).
13. Ono, T., Kasamatsu, A., Oka, S. & Moss, J. The 39-kDa poly(ADP-ribose) glycohydrolase ARH3 hydrolyzes O-acetyl-ADP-ribose, a product of the Sir2 family of acetyl-histone deacetylases. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 16687–16691 (2006).
14. Nottbohm, A. C., Dothager, R. S., Putt, K. S., Hoyt, M. T. & Hergenrother, P. J. A Colorimetric Substrate for Poly(ADP-Ribose) Polymerase-1, VPARP, and Tankyrase-1. *Angew. Chem. Int. Ed.* **46,** 2066–2069 (2007).
15. Pitsch, S. An Efficient Synthesis of Enantiomeric Ribonucleic Acids from D-Glucose. *Helv. Chim. Acta* **80,** 2286–2314 (1997).
16. Messner, S. *et al.* PARP1 ADP-ribosylates lysine residues of the core histone tails. *Nucleic Acids Res.* **38,** 6350–6362 (2010).
17. Tao, Z., Gao, P. & Liu, H.-W. Identification of the ADP-Ribosylation Sites in the PARP-1 Automodification Domain: Analysis and Implications. *J. Am. Chem. Soc.* **131,** 14258–14260 (2009).
18. Moyle, P. M. & Muir, T. W. Method for the Synthesis of Mono-ADP-ribose Conjugated Peptides. *J. Am. Chem. Soc.* **132,** 15878–15880 (2010).
19. Chapman, J. D., Gagné, J.-P., Poirier, G. G. & Goodlett, D. R. Mapping PARP-1 Auto-ADP-ribosylation Sites by Liquid Chromatography–Tandem Mass Spectrometry. *J. Proteome Res.* **12,** 1868–1880 (2013).
20. Kun, E., Kirsten, E., Mendeleyev, J. & Ordahl, C. P. Regulation of the Enzymatic Catalysis of Poly(ADP-ribose) Polymerase by dsDNA, Polyamines, Mg^{2+}, Ca^{2+}, Histones H1 and H3, and ATP. *Biochemistry* **43,** 210–216 (2003).
21. Kiehlbauch, C. C. C., Aboul-Ela, N., Jacobson, E. L. E., Ringer, D. P. D. & Jacobson, M. K. M. High resolution fractionation and characterization of ADP-ribose polymers. *Anal. Biochem.* **208,** 26–34 (1993).
22. Dunstan, M. S. *et al.* Structure and mechanism of a canonical poly(ADP-ribose) glycohydrolase. *Nat. Commun.* **3,** 878–6 (2012).
23. Barkauskaite, E. *et al.* Visualization of poly(ADP-ribose) bound to PARG reveals inherent balance between exo- and endo-glycohydrolase activities. *Nat. Commun.* **4,** 1–8 (2013).

Chapter 5. Mode of Action Studies of the Small Molecule Raptinal

Portions of this chapter are reprinted with permission from: Palchaudhuri, R.; Lambrecht, M.J.; Botham, R.; Partlow, K.; van Ham, T.J.; Putt, K.S.; Nguyen, L.T; Kim, S.; Peterson, R.T.; Fan, T.; Hergenrother, P.J.; A Small Molecule that Induces Intrinsic Pathway Apoptosis with Unparalleled Speed. *Cell Rep.* **2015**, 13, 2027-2036. The contributions of others are mentioned throughout.

5.1 Introduction to Raptinal

The induction of apoptosis is a major mechanism through which chemotherapeutics kill cancer cells[1] and the misregulation of apoptosis is a hallmark of cancer[2]. As such, new molecules that induce apoptosis often present promising leads for the treatment of cancers. Also of significant importance is the discovery of novel mechanisms of action and phenotypes through which to trigger apoptosis. Such discoveries could potentially lead to new classes of chemotherapeutics or inform understanding of apoptotic processes.



Raptinal (5-1)

Figure 5.1. The chemical structure of Raptinal, an unusually fast inducer of apoptosis.

The small molecule Raptinal (Figure 5.1) was discovered in a high throughput screen of the UIUC Heritage Library by Dr. Karson Putt and shown by Dr. Rahul Palchaudhuri to be an unusually rapid inducer of apoptotic cell death. Still, how exactly Raptinal induces apoptosis was

not (and even today is not) entirely clear. The experiments herein were performed to elucidate the mode of action of Raptinal.

5.2 Raptinal Exists Primarily as a Hydrate in Aqueous Solution

While aldehydes and ketones are in equilibrium with their corresponding gem-diol (hydrate) species in aqueous solution, generally the equilibrium favors the non-hydrated structure in the absence of strong electron withdrawing substituents[3]. However, it was imagined that the dialdehyde structure of Raptinal might lend itself to hydration as is seen for structurally similar dialdehyde molecules such as o-phthalaldehyde[4]. To test this hypothesis, Raptinal was incubated in solutions containing d_6-DMSO and increasing concentrations of D_2O. It was found that Raptinal converted to a new species by ^1H NMR spectroscopy (Figure 5.2) and that this conversion was quite rapid at high concentrations of D_2O (Figure 5.3). Two-dimensional NMR spectroscopy was used to confirm that the new species formed were the hydrates of Raptinal and enabled the assignment the resonances in figure 5.2. Particularly striking is the disappearance of the aldehyde resonance (A, figure 5.2) and the corresponding appearance of the hemiacetal resonance (9, figure 5.2) in its place. Additionally, due to the *cis* or *trans* disposition of the -OH groups on the hydrate, two hydrated isomers are formed and the aryl rings on the fluorene of each isomer become magnetically non-equivalent.

The NMR-based approach to monitoring Raptinal hydration only allowed for D_2O to comprise 25% of the solution, above which point reasonable NMR signals could not observed due to the lack of solubility and the millimolar concentrations of compound necessary for detection by NMR spectroscopy. However, for cell culture experiments where water contents are much higher (generally 99% H_2O and 1% DMSO with Raptinal concentrations of 10 μM)

hydration is likely even faster and the hydrate is likely the predominant species in solution. The ability of the pre-formed hydrate to induce rapid apoptotic cell death was assessed and compared to non-hydrated Raptinal which was added to cells as a DMSO solution. It was found that both treatments caused > 80% cell death at 2 h as assessed by flow cytometry and that they had identical 24 h IC_{50} values (840 nM) in U-937 cells as determined by the MTS assay.

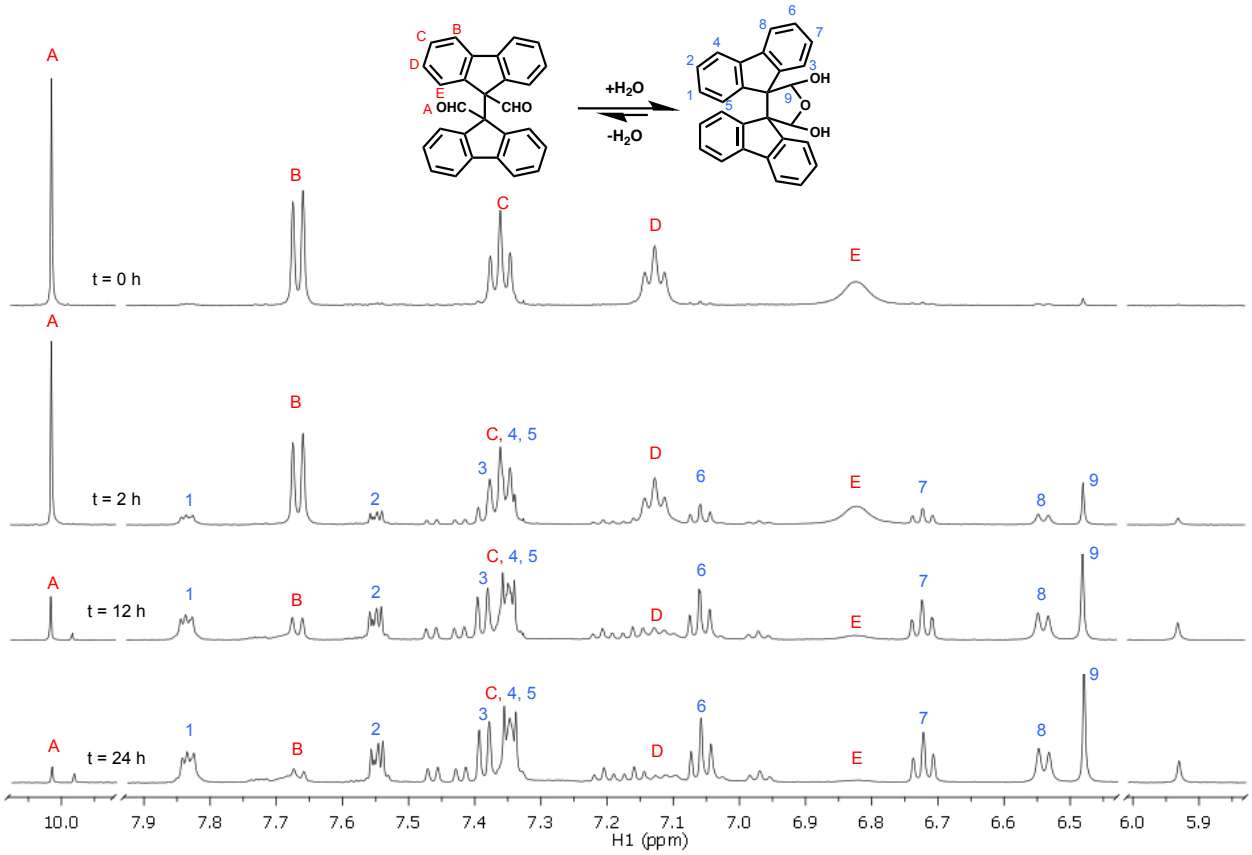


Figure 5.2. Hydration of Raptinal as seen by ^{1}H NMR Spectroscopy. Only the major hydrated isomer is annotated for clarity of the figure, though the minor isomer can be seen.

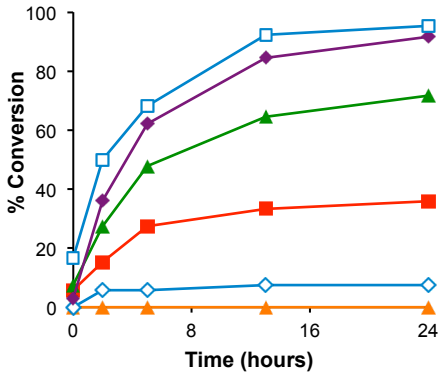


Figure 5.3. Conversion of Raptinal to hydrate as measured by ^{1}H NMR spectroscopy. Quantification was performed by comparison of integration of signal E of the dialdehyde form to signal 7 of the hydrate form (Figure 5.2) as these were cleanly resolved from other resonances.

5.3 The Synthesis and Biological Evaluation of Analogs of Raptinal

With the desire to perform traditional target identification studies[5] of Raptinal via a compound-tagged approach, analogs were synthesized (Figure 5.4) to develop a structure-activity relationship profile. Raptinal itself was synthesized via a known[6] two-step procedure on multigram scale. Additionally, Raptinal could be used as a substrate for reduction to form **5-3** or to form oxime **5-4**. Treatment of the olefin **5-5** with sodium metal with trapping of carbon dioxide afforded the diacid **5-6** from which the anhydride **5-7** and methyl ester **5-8** were synthesized. Synthesis of brominated (**5-9**) and chlorinated (**5-10**) analogs was performed by the method used to synthesize Raptinal. Also, ketone (**5-11**) and trifluoromethyl ketone (**5-12**) analogs were prepared.

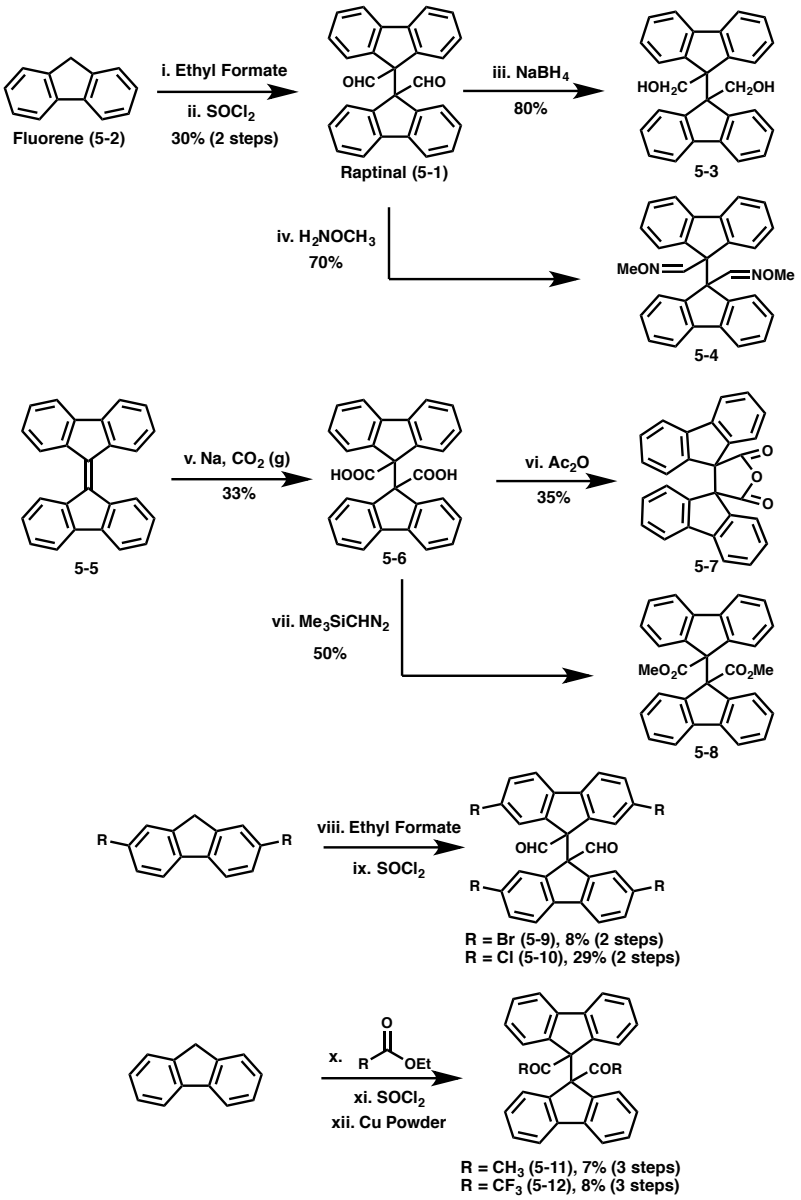


Figure 5.4. Synthesis Raptinal (**5-1**) and analogs. i. Ethyl formate (2.5 eq), KOMe (2 eq), Et_2O, reflux, 30 min. ii. $SOCl_2$, Et_2O, reflux, 3 h, 30% (2 steps). iii. $NaBH_4$ (8.5 eq), EtOH, rt, 45 min, 80%. iv. $H_2NOCH_3{\cdot}HCl$ (2.5 eq), pyridine, rt, 44 h, 70%. v. Na (4.3 eq), Et_2O, rt 48 h, then: CO_2, rt, 15 min, 33%. vi. Ac_2O, 70°C, 2 h, 35%. vii. Me_3SiCHN_2 (3 eq), MeOH, rt, 2 h, 50%. viii. KOMe (2 eq), ethyl formate (2.5 eq), Et_2O, reflux, 1.5 h. ix. $SOCl_2$, Et_2O, reflux 17 h, 8% (2 steps for **5-9**), 29% (2 steps for **5-10**). x. KOtBu (2 eq), RCO_2Et (1.5 eq), THF, relux, 4 h (56% for R = CH_3, 64% for R = CF_3). xi. $SOCl_2$, reflux, 24 h (34% for R = CH_3). xii. Cu powder, benzene, reflux, 20 h (38% for R = CH_3, 13% for R = CF_3 for 2 steps).

Upon accessing these compounds, they were assessed by Rachel Botham for their ability to induce cell death in U-937 cells (Table 5.1). Cell death assays were performed at both the 2 h

and 24 h time points in an effort to determine the potency of these analogs relative to Raptinal, but also to observe if they retain the rapid cell death phenotype of the parent molecule. These experiments reveal that few of these compounds display any measurable activity against U-937 cells in the 24 h assay. Those compounds that do, specifically **5-9** and **5-10**, show greatly reduced activity at 2 h relative to Raptinal. From this data we concluded that the aldehyde moiety of Raptinal is absolutely essential for biological activity, and its removal leads to compounds that are non-cytotoxic. Compounds **5-14**, **5-15**, and **5-16** show that the presence of the 1,4-dialdehyde is not sufficient to induce cytotoxicity, as these compounds were unable to induce cell death at similar concentrations as Raptinal.

As most modifications slowed or abolished the phenotype of Raptinal, we concluded that derivative synthesis and pull-down experiments would be unlikely to further elucidate the mechanism of the compound. A further complication of derivative synthesis and pull-down was the challenges associated with synthesizing dialdehyde Raptinal derivatives. Such compounds were found to be unstable to silica column chromatography. Additionally, byproducts from synthesis were challenging to remove by chromatography on florisil, a medium on which they were stable. Purification by reverse phase chromatography rarely produced clean separation due to the hydration phenomenon described previously.

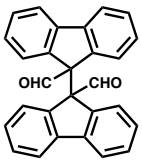
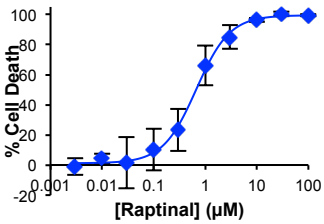
Compound	24 h IC_{50} (μM)	24 h IC_{50} Curve	% Cell Death at 2 h (10 μM treatment)
 **Raptinal (5-1)**	**0.7 ± 0.3**		**84.9 ± 17.7**

Table 5.1. Evaluation of cell death induced by Raptinal derivatives at 2 h and 24 h in U-937 cells. Experiments performed by Rachel Botham.

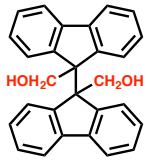
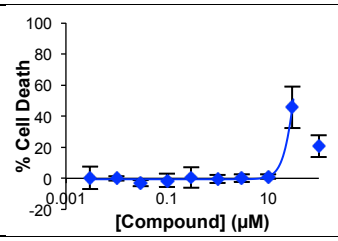
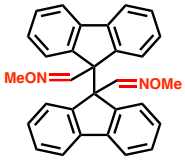
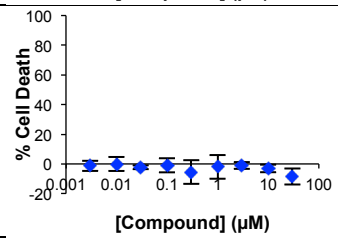
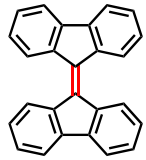
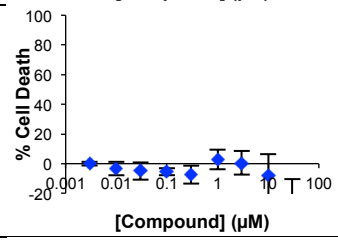
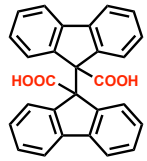
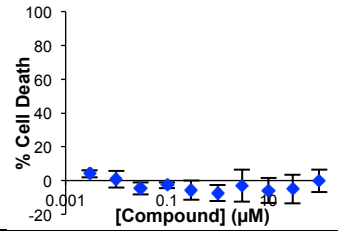
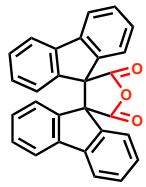
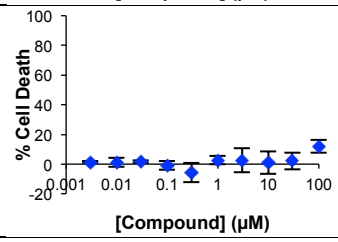
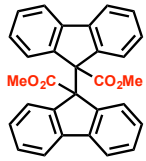
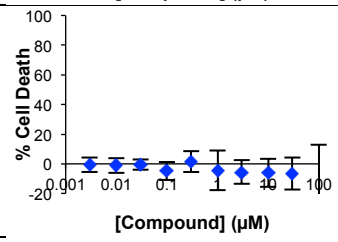
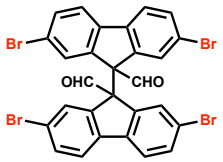
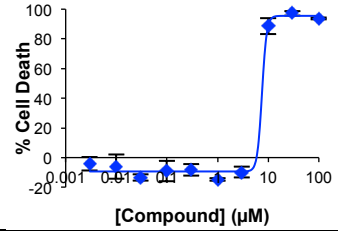
 **5-3**	**33 ± 7**		**N.D.**
 **5-4**	> 100		N.D.
 **5-5**	> 100		N.D.
 **5-6**	> 100		N.D.
 **5-7**	> 100		N.D.
 **5-8**	> 100		N.D.
 **5-9**	6.5 ± 1.6		6.3 ± 0.8

Table 5.1. (continued)

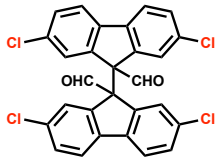
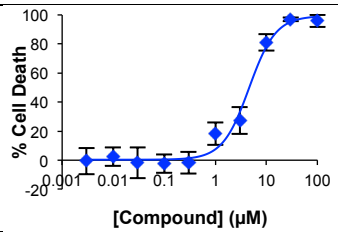
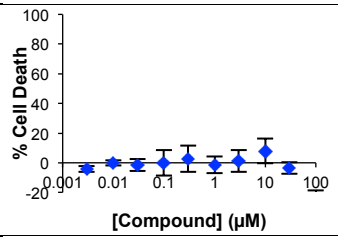
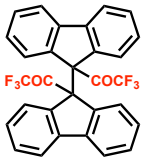
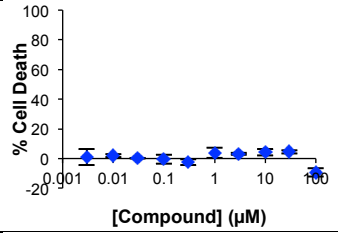
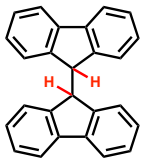
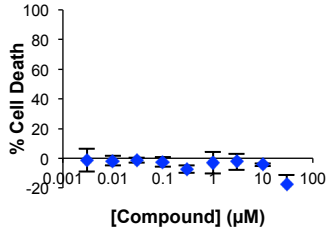
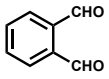
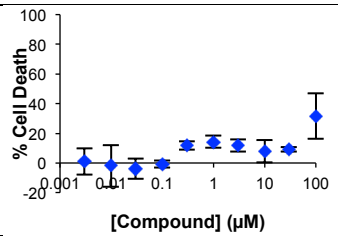
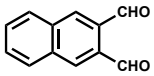
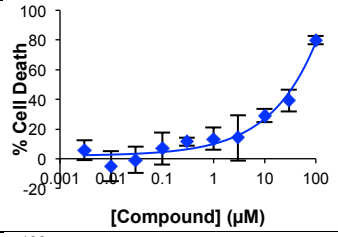
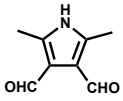
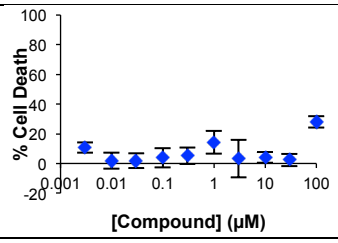
 **5-10**	**4.6 ± 0.9**		**7.6 ± 1.9**
 **5-11**	> 100		N.D.
 **5-12**	> 100		N.D.
 **5-13**	> 100		N.D.
 **5-14**	> 100		N.D.
 **5-15**	38 ± 12		N.D.
 **5-16**	> 100		N.D.

Table 5.1. (continued)

5.4 RNAi as a Method for Target Identification and Validation of Raptinal

Whole genome shRNA screens have confirmed protein targets of small molecules with known modes of action[7]. While this approach had not previously been reported in *de novo* target identification of small molecules, Dr. Rahul Palchaudhuri performed such an experiment for Raptinal with the hope that this method would uncover its mode of action. For this work, cells were transfected with shRNA and then treated with Raptinal; the assumption is that cells that survive treatment are able to do so because of knockdown of a gene essential in Raptinal-induced cell death. Upon performing this shRNA screen for Raptinal, 142 genes with >40 fold enrichment for Raptinal and 45 that were enriched >5 fold with two or more transcripts per gene were identified. In addition to these 187 (142 single transcirpt + 45 two or more transcripts) an additional 58 genes known to be associated with apoptosis were selected for follow-up (for a total of 142 + 45 + 58 = 245). For follow-up, siRNAs were selected (single siRNA per gene) and their ability to prevent Raptinal-induced caspase-3 activity in a cell-lysate caspase-3/-7 activity assay using a fluorescent substrate was determined. Dr. Palchaudhuri performed the initial rescreen of siRNAs and found 44 "rehits," or siRNAs that resulted in <40% of maximal caspase activation using the fluorescence assay.

With 44 rehits protecting from Raptinal-induced cell death to varying degrees, we sought to understand which of these RNAs were specifically protecting from Raptinal and which were more generally protective from apoptosis, but were not necessarily specific to Raptinal's mechanism. To interrogate this, a co-screen was conducted of these rehits compared to staurosporine (STS) and doxorubicin, two inducers of apoptosis that we expected to mechanistically induce apoptosis in different ways than Raptinal (Figure 5.5). Very few RNAs were found that were selective for Raptinal over both STS and doxorubicin, suggesting that

many of the siRNAs in the collection were protecting generally from apoptosis and less from a specific aspect of Raptinal's mode of action. Additionally, we were disappointed to find that the magnitude of protection for most siRNAs was small compared to the robust protection seen by those for apoptotic proteins such as caspase-3, diablo, and APAF-1. If a given protein was directly involved in Raptinal's mode of action, we would expect its knockdown by siRNA would nearly quantitatively protect from induction of caspase-3 activity. The large degree of protection from RNAs for apoptotic proteins such as caspase-3, APAF-1, and diablo was further support for Raptinal-induced apoptosis being highly consistent with the intrinsic pathway.

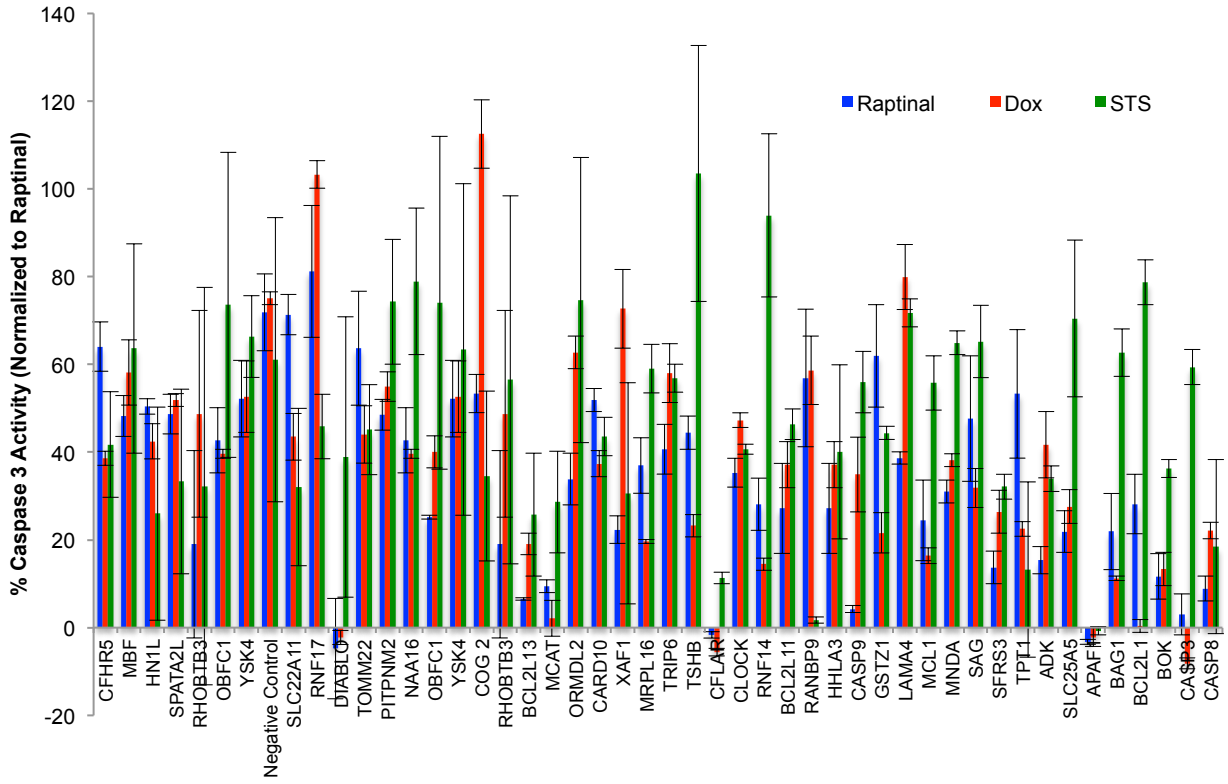


Figure 5.5. Percent caspase-3 activity after treatment with Raptinal, doxorubicin (Dox), or staurosporine (STS) for 44 follow-up siRNAs in MIA PaCa-2 cells. These experiments were performed with assistance from Dr. Kathy Partlow.

5.5 Small Molecule Cytoprotectants to Inform Raptinal's Mode of Action

In addition to RNAi as a method through which to specifically target genes (and therefore proteins) to find targets important in Raptinal's mode of action, small molecules with known biological activity were tested for their ability protect from Raptinal-induced cell death (Figure 5.6). The strongest protection was seen from agents that affect the electron transport chain. For instance, inhibitors of complex I (rotenone), complex III (anitmycin A), complex IV (sodium azide and potassium cyanide) and ATP synthase (oligomycin) all provided high levels of protection as did a mitochondrial respiration uncoupler (FCCP). These results are consistent with previous reports of oligomycin protecting from cytochrome *c* release in apoptosis[8] and the importance of the mitochondrial proton pump in apoptosis[9]. While the exact mechanisms of how these reagents affecting the mitochondrial electron transport chain protect from apoptosis is not entirely clear, it is reasonable given the important role of cytochrome *c* in both processes.

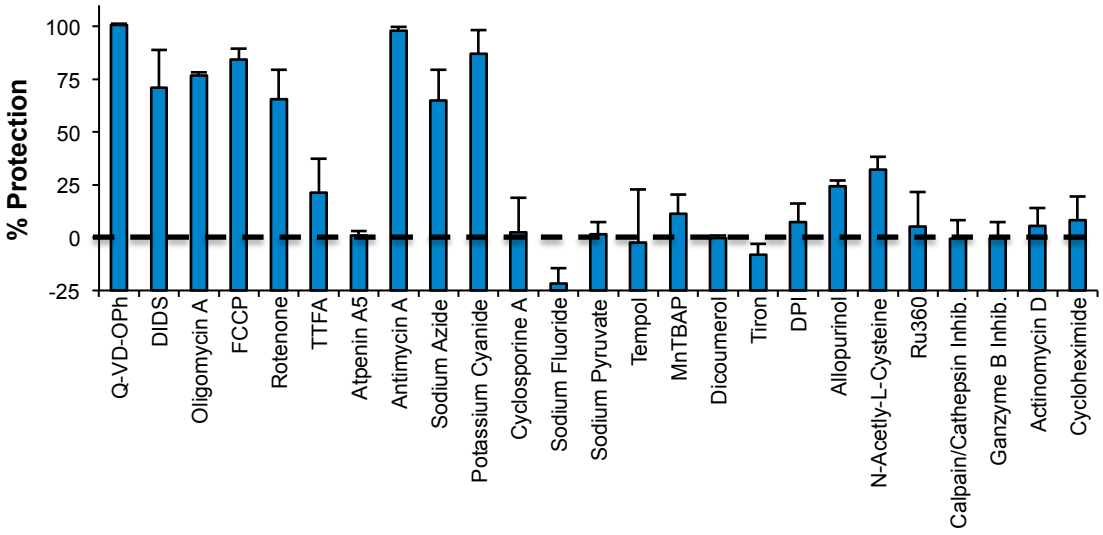


Figure 5.6. Raptinal-induced cell death can be protected by treatment with various small molecule cytoprotectants. All experiments were performed with 2 h pretreatment with cytoprotectants at optimized concentrations and with 2 h treatment of Raptinal at 10 μM in U-937 cells. Cell death was assessed by flow cytometry. Data for certain compounds was contributed by Dr. Rahul Palchuldhuri and Rachel Botham.

5.6. Conclusions and Outlook on Raptinal

The small molecule Raptinal is especially unique among small molecules given its ability to induce rapid and clean apoptosis across many cell lines. While the exact mechanism by which it triggers induction of the apoptotic cascade is not entirely clear, the data in figures 5.5 and 5.6 clearly supports that it is entirely consistent with the intrinsic (mitochondrial) pathway of apoptosis. Given the flat SAR of Raptinal as described in section 5.3, genetic approaches to target identification similar to the shRNA screen/siRNA follow-up will likely be the most fruitful going forward. The challenge with such approaches is distinguishing candidate targets from proteins that are identified as hits because they are generally protective from apoptosis processes rather than specific to a given molecule. In the future, an experiment such as the shRNA screen should be done with co-screening with mechanistically similar molecules to Raptinal such as STS. While Dr. Palchaudhuri did some triaging and comparison of Raptinal shRNAs with those for DNQ, the mechanistic dissimilarities between the two compounds (DNQ induces cell death through a ROS-based mechanism inconsistent with apotosis) makes such analyses less useful than had they been done for agents that directly induce apoptosis.

5.7 Experimental

Cell culture conditions

All cells were grown in RPMI 1640, DMEM or EMEM media supplemented with 10% FBS, 1% penicillin-streptomycin and incubated at 37 ºC in 5% CO_2, 95% humidity incubator.

MTS assay for suspension cells

Serial dilutions of compound in 100% DMSO were added in triplicate (2 μL to each well) to empty wells of a 96-well plate. Suspension cells (HL-60, U-937, SKW 6.4, Jurkat WT, Jurkat CASP8$^{-/-}$, Jurkat FADD$^{-/-}$, Jurkat Bcl-2) cells in RPMI 1640 media were added to 96-well plates (198 μL containing 4 x 10^4 cells) and the cells incubated for 24 hours. A solution containing the soluble tetrazolium salt ((3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt; MTS) and the electron coupling reagent, phenazine methosulfate (PMS) was prepared according to the manufacturer's instructions (Promega) and 20 μL added to each well. The plates were incubated at 37 ºC for 15-45 min and the absorbance at 490 nm was measured using a SpectraMax Plus 384 well plate reader (Molecular Devices, Sunnyvale CA). The mean IC_{50} values and standard deviations were determined from three independent experiments.

Protection assays using small molecule inhibitors

For protection assays, U-937 cells (0.5 x 10^6 cells/mL) were pretreated with the prospective protective agents for 2 hours at the following concentrations: Sodium fluoride (1 mM), cyclosporine A (10 μM), oligomycin A (10 μM), FCCP (10 μM), cycloheximide (10 μM), TTFA (1 mM), atpenin A5 (1 μM), actinomycin D (2 μM), rotenone (200 μM), potassium cyanide (1 mM), sodium azide (1 mM) antimycin A (100 μM), sodium pyruvate (5 mM), tempol (400 μM), granzyme B inhibitor I (50 μM), calpain inhibitor I (50 μM), MnTBAP (20 μM), DPI (10 μM), allopurinol (1 mM), dicoumarol (50 μM), tiron (10 mM), Ru360 (10 μM), NAC (10 mM), DIDS (2 mM), and Q-VD-OPh (50 μM). The cells were then co-treated with **Raptinal** at

10 μM for 2 hours prior to analysis by propidium iodide/FITC-annexin V staining and flow cytometry.

Transfection and screening of siRNA constructs

All RNA interference reagents were purchased from Life Technologies (Carlsbad, CA). The optimal transfection conditions (i.e., siRNA concentration, time, cell number, etc.) were determined using the KDalert GAPDH Assay Kit according to the manufacturer's instructions. Human Silencer Select siRNA was purchased for gene targets of interest and diluted in DEPC water. The siRNA (5 nM final conc.) was complexed in OptiMEM media with RNAiMAX lipofectamine (0.3 μL lipofectamine/pmol siRNA) for 5 min in the well prior to addition of cells. MIA PaCa-2 cells were trypsinized, washed in PBS, resuspended at 50 cells/μL in OptiMEM containing 4% FBS, added to the well (100 μL/96-well or 2.5 mL/6-well), and grown for 4 days. Silencer® Negative Control siRNA was used in control wells. After silencing, fresh OptiMEM with 4% FBS was supplied to the cells (half immediately and half at time of compound addition).

Chemistry

General

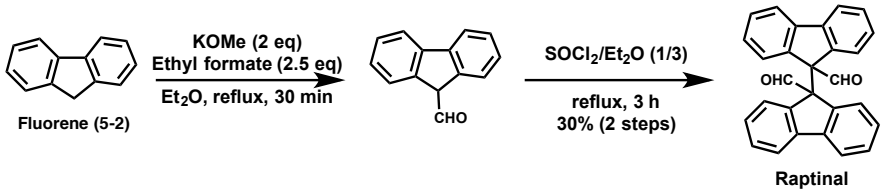
Compounds **5-5**, **5-13**, **5-14** and **5-15** are commercially available, were purchased and were used as received. All reactions were run in flame or oven dried glassware under an atmosphere of dry nitrogen unless otherwise noted. Acetonitrile, tetrahydrofuran, methanol and methylene chloride used in reactions were obtained from a solvent dispensing system. Diethyl ether was distilled from sodium metal. 4 Å molecular sieves were dried at 200 °C on high

vacuum overnight. Pyridine was distilled from CaH_2 and stored on 4 Å molecular sieves. All other reagents were of standard commercial purity and were used as received.

Analytical thin-layer chromatography was performed on EMD Merck silica gel plates with F254 indicator. Plates were visualized with UV light (254 nm) or staining with p-anisaldehyde. Silica gel for column chromatography was purchased from Sorbent Technologies (40-75 μm particle size).

Unless otherwise indicated, ^{1}H, ^{13}C, ^{19}F, and ^{31}P NMR spectra were recorded at 500, 125, 470 and 203 MHz, respectively. ^{1}H and ^{13}C NMR spectra were referenced to tetramethylsilane or the residual solvent peak. ^{19}F NMR spectra were referenced using C_6F_6 as an internal standard (-164.9 ppm). Chemical shifts are reported in ppm and multiplicities are reported as s (singlet), d (doublet), t (triplet), q (quartet), p (pentet), h (hextet), hep (heptet), m (multiplet), and b (broad). Mass spectrometry analysis was performed by the University of Illinois Mass Spectrometry Center.

Procedures for Chemical Synthesis



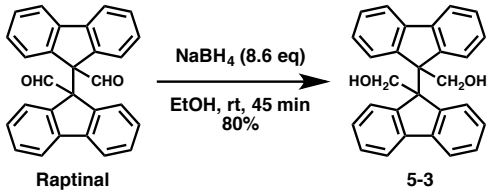
Raptinal (5-1): A modified version of the procedure of Curtin and co-workers was followed[6]. To a solution of fluorene (**5-2**) (4.4 g, 26.4 mmol, 1 eq) and potassium methoxide (3.7 g, 52.8 mmol, 2 eq) in diethyl ether (50 mL) was added ethyl formate (5.3 mL, 66.0 mmol, 2.5 eq) and the reaction was heated at reflux for 30 min. The reaction was poured into 1M KOH, and the organic layer removed. The aqueous layer was acidified with concentrated HCl until the solution

tested acidic to pH paper (at which time a white precipitate was formed). The solution was extracted with diethyl ether, dried through Na_2SO_4, and concentrated. The crude material was dissolved in diethyl ether (25 mL), thionyl chloride (8 mL) was added and the solution was heated to reflux for 3 h. The reaction was quenched by careful addition to ice water. The solution was extracted three times with methylene chloride, washed with water and saturated aqueous NaCl, dried through Na_2SO_4, and concentrated. The crude material was recrystallized from acetic acid, the crystals were collected and washed with water until the washing solution tested neutral to pH paper, affording **Raptinal (5-1)** as white crystals (1.52 g, 30%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 9.89 (s, 2H), 7.53 (dt, *J* = 7.6, 1.0 Hz, 4H), 7.33 (td, *J* = 7.5, 1.1 Hz, 4H), 7.16 – 7.06 (m, 4H), 6.97 (bd, *J* = 7.8 Hz, 4H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 197.6, 142.5, 139.6, 129.2, 127.2, 126.8, 119.9, 71.2.

HRMS (ESI): m/z 387.1381, [calculated for $C_{28}H_{19}O_2$ $(M+H)^+$: 387.1385]

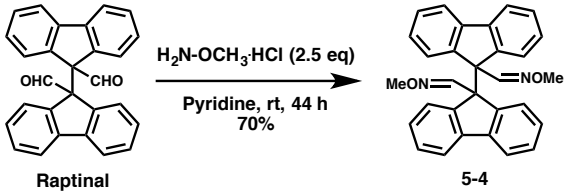


Compound 5-3: To a solution of **Raptinal (5-1)** (32 mg, 0.0828 mmol, 1 eq) in EtOH (0.8 mL) was added $NaBH_4$ (27 mg, 0.714 mmol, 8.6 eq) and the reaction was stirred for 45 min. The reaction was quenched by careful addition of 1M HCl. The reaction was extracted three times with $CHCl_3$, dried through Na_2SO_4, and concentrated. The compound was purified by silica column chromatography to give compound **5-3** (26 mg, 80%) as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.54 (d, *J* = 7.5 Hz, 4H), 7.27 (t, *J* = 7.5 Hz, 4H), 7.07 (t, *J* = 7.5 Hz, 4H), 6.96 (bs, 4H), 4.10 (s, 4H), 3.45 (bs, 2H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 145.3, 141.3, 127.9, 126.6, 125.1, 119.7, 67.0, 60.4.

HRMS (ESI): m/z 413.1510, [calculated for $C_{28}H_{22}O_2Na$ $(M+Na)^+$: 413.1512]



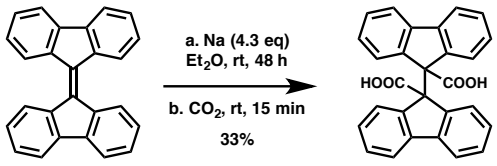
Raptinal 5-4

Compound 5-4: To a solution of **Raptinal (5-1)** (32 mg, 0.083 mmol, 1 eq) in pyridine (0.8 mL) was added methoxyamine hydrochloride (17 mg, 0.207 mmol, 2.5 eq) and the reaction was stirred at room temperature for 44 h. The reaction was diluted with chloroform and washed twice with 1M HCl. The organic layer was dried through Na_2SO_4, concentrated, and purified by silica column chromatography to afford compound **5-4** (26 mg, 70%) as a white foam.

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.35 (s, 2H), 7.45 (d, J = 7.6 Hz, 4H), 7.25 (t, J = 7.7 Hz, 4H), 7.06 (t, J = 7.0 Hz, 4H), 6.83 (s, 4H), 3.88 (s, 6H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 150.1, 143.0, 141.2, 128.2, 126.3, 126.3, 119.4, 62.1, 60.8.

HRMS (ESI): 445.1912 [calculated for $C_{30}H_{25}N_2O_2$ $(M+H)^+$ 445.1916]



5-5 5-6

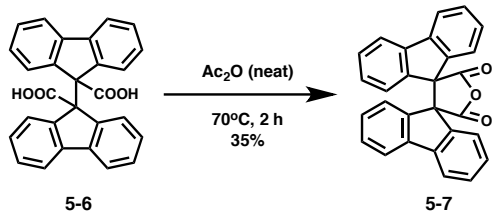
Compound 5-6: To a solution of compound **5-5** (100 mg, 0.304 mmol, 1 eq) in diethyl ether (6 mL) was added sodium metal (30 mg, 1.30 mmol, 4.3 eq). The solution was stirred at room temperature for 48 h, at which point CO_2 (g) was bubbled through the solution. The crude reaction was poured into 2.5% NaOH, and washed once with diethyl ether. The aqueous layer

was acidified until complete precipitation at which point it was extracted three times with methylene chloride and once with diethyl ether. The organic extracts were combined, dried through Na_2SO_4, concentrated and purified by silica column chromatography (with 1% formic acid) to afford compound **5-6** (42 mg, 33%).

^{1}H NMR (500 MHz, CD_3OD) δ 7.42 (d, *J* = 7.5 Hz, 4H), 7.23 (t, *J* = 7.5 Hz, 4H), 7.08 (bs, 4H), 7.00 (t, *J* = 7.5 Hz, 4H)

^{13}C NMR (125 MHz, CD_3OD) δ 174.7, 143.9, 143.1, 129.2, 128.6, 126.9, 119.8, 67.7.

HRMS (ESI): m/z 417.1125, [calculated for $C_{28}H_{17}O_4$ $(M-H)^-$: 417.1127]



5-6 5-7

Compound 5-7: To a reaction vial was added compound **5-6** (50 mg, 0.12 mmol, 1 eq) and acetic anhydride (2 mL). The reaction was heated to 70 °C for 2 h, after which the solvent was evaporated. The compound was purified by silica column chromatography to give 16.6 mg (35%) of compound **5-7**.

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.58 (d, *J* = 8 Hz, 4H), 7.39 (bt, *J* = 8 Hz, 4H), 7.20 (bs, 8H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 169.9, 141.7, 138.4, 130.3, 127.8, 124.7, 120.9, 68.1

HRMS (ESI): m/z 423.1008, [calculated for $C_{28}H_{16}O_3Na$ $(M+Na)^+$: 423.0997]

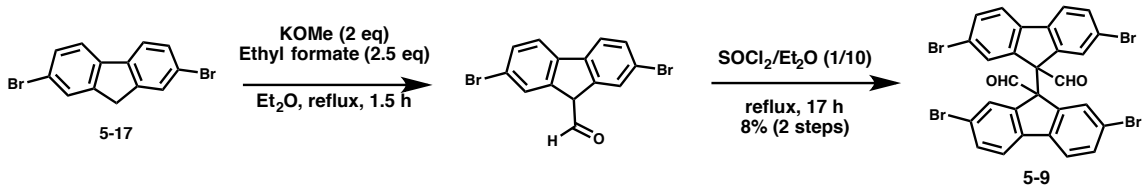
5-6 5-8

Compound 5-8: To a room temperature solution of compound **5-6** (10 mg, 0.0239 mmol, 1 eq) in methanol (0.3 mL) was added TMS-diazomethane (0.036 mL of a 2M solution in diethyl ether, 0.717 mmol, 3 eq) at room temperature. The reaction was stirred at room temperature for 2 h. The reaction was concentrated and purified by silica column chromatography to afford compound **5-8** (5.3 mg, 50%) as a white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.39 (dd, *J* = 7.6, 1.2 Hz, 4H), 7.25 (dd, *J* = 7.8, 6.6 Hz, 4H), 7.08 – 7.01 (m, 4H), 6.97 (bd, *J* = 7.9 Hz, 4H), 3.77 (s, 6H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 171.8, 141.7, 141.5, 128.6, 127.4, 126.7, 126.3, 119.0, 66.4, 52.9.

HRMS (ESI) 447.1603 [calculated for $C_{30}H_{23}O_4$ $(M+H)^+$ 447.1596]

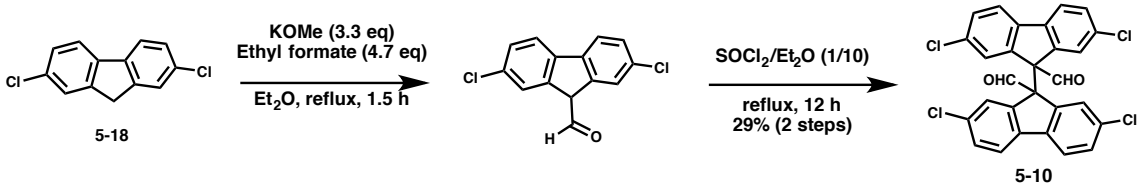


Compound 5-9: To a solution of 2,7-dibromofluorene (**5-17**) (500 mg, 1.5 mmol, 1 eq) and potassium methoxide (210 mg, 3.0 mmol, 2 eq) in diethyl ether (10 mL) was added ethyl formate (0.3 mL, 3.75 mmol, 2.5 eq) and the reaction was heated to reflux for 1.5 h. The reaction was poured into water and washed with diethyl ether. The aqueous layer was acidified with concentrated HCl until precipitation was observed. The aqueous layer was extracted twice with diethyl ether, washed once with water, dried through Na_2SO_4 and concentrated. To the crude material was added diethyl ether (10 mL) and thionyl chloride (1 mL) and the reaction was heated at reflux for 17 h. The product was precipitated and washed with diethyl ether to afford compound **5-9** (42 mg, 8%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 9.80 (s, 2H), 7.49 (d, *J* = 6.6 Hz, 4H), 7.37 (d, *J* = 8.1 Hz, 4H), 7.03 (bs, 4H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 196.2, 141.0, 140.8, 133.4, 130.6, 121.9, 121.8, 71.5.

HRMS (ESI) 724.7595, [calculated for $C_{28}H_{14}O_2Br_4Na$ $(M+Na)^+$ 724.7581]

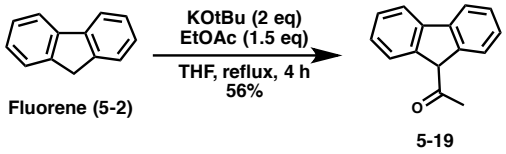


Compound 5-10: To a solution of 2,7-dichlorofluorene (**5-18**) (1.0 g, 4.25 mmol, 1 eq) and potassium methoxide (1.0 g, 14.2 mmol, 3.3 eq) in diethyl ether (30 mL) was added ethyl formate (1.6 mL, 19.8 mmol, 4.7 eq) and the reaction was heated to reflux for 2.5 h. The reaction was poured into water and extracted twice with hexane. The aqueous layer was acidified with concentrated HCl until precipitation was observed. The aqueous layer was extracted three times with ethyl acetate, the combined organic extracts were washed with saturated aqueous NaCl, dried through Na_2SO_4 and concentrated. The crude product was dissolved in diethyl ether (5 mL) and thionyl chloride (0.62 mL) was added. The reaction was heated at reflux for 12 h. The precipitate was collected and washed with diethyl ether to give compound **5-10** (76 mg, 29%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 9.82 (s, 2H), 7.44 (d, *J* = 8.1 Hz, 4H), 7.36 (d, *J* = 8.2 Hz, 4H), 6.91 bs, 4H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 196.2, 140.9, 140.4, 133.9, 130.5, 127.7, 121.5, 71.3.

HRMS (ESI) 522.9849 [calculated for $C_{28}H_{15}O_2Cl_4$ $(M+H)^+$ 522.9826]

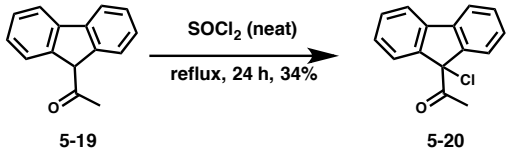


Compound 5-19: The procedure of Borowiecki and co-workers was followed and the product matches the known compound [10]. To a 50 mL round-bottom flask was added fluorene (**5-2**) (1.0 g, 6.02 mmol), and potassium t-butoxide (1.01 g, 9.03 mmol, 1.5 eq). The solids were dissolved in THF (20 mL) and ethyl acetate (0.89 mL, 9.03 mmol, 1.5 eq) was added dropwise to the solution. The reaction was heated at reflux under nitrogen for 4 h. The reaction mixture was poured into a saturated solution of NH_4Cl and which was extracted three times with diethyl ether. The combined ether extracts were washed with a saturated NaCl solution, dried over Na_2SO_4, and concentrated *in vacuo*. The compound was purified by silicia column chromatography to give compound **5-19** as a yellow solid (699 mg, 56%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.82 (d, J=8 Hz, 2H), 7.51 (d, J=7.5 Hz, 2H), 7.46 (t J=7.5 Hz, 2H), 7.35 (t, J=7.5 Hz, 2H), 4.80 (s, 1H), 1.62 (s, 3H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 207.1, 142.6, 141.6, 128.9, 128.2, 125.6, 120.9, 64.0, 25.6

HRMS (ESI): 209.0970 [calculated mass for $C_{15}H_{13}O$ $(M+H)^+$ 209.0966]



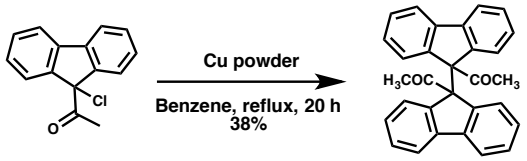
Compound 5-20: A modified version of the procedure of Greenhow and co-workers was followed[11]. To a 10 mL round bottom flask containing compound **5-19** (1.0 g, 4.8 mmol) was added thionyl chloride (4 mL). The reaction was heated at reflux for 24 h. Following removal of thionyl chloride, the product was purified by silica column chromatography (10:1 hexane/ethyl

acetate) and recrystallized from ethanol, affording compound **5-20** (395 mg, 34%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.76 (d, J=7.2 Hz, 2H), 7.51 (d, J=7.2 Hz, 2H), 7.49 (td, J=7.6 Hz, 0.8 Hz, 2H), 7.39 (td, J=7.6 Hz, 1.2 Hz, 2H), 1.770 (s, 3H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 198.8, 144.2, 141.0, 130.8, 129.4, 125.6, 121.3, 110.0, 24.9

HRMS (ESI): 265.0399 [calculated mass for $C_{15}H_{11}OClNa$ $(M+Na)^+$ 265.0396]



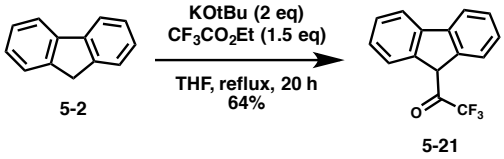
5-20 5-11

Compound 5-11: A modified version of the procedure of Greenhow and co-workers was followed [11]. To a 25 mL round bottom flask containing **5-20** (169 mg, 0.69 mmol, 1 eq) was added 274 mg of copper powder and 15 mL of benzene. The reaction was heated at reflux for 20 h, the copper was filtered and the solvent was evaporated. The product was purified by silicia column chromatography (10:1 hexane/ethyl acetate) to give compound **5-11** (55.3 mg, 38%). Note: NMR analysis of the product at room temperature resulted in very poor NMR signal, presumably due to limited rotation about the C9-C9' single bond. NMR spectroscopic analysis was therefore carried out at -40 °C.

^{1}H NMR (500 MHz, $CDCl_3$, -40^{o}C) δ 8.24 (d, *J* = 7.5 Hz, 2H), 7.56 (d, *J* = 7.3 Hz, 2H), 7.53 – 7.41 (m, 4H), 7.29 (d, *J* = 7.7 Hz, 2H), 7.04 (t, *J* = 7.5 Hz, 2H), 6.63 (t, *J* = 7.5 Hz, 2H), 5.86 (d, *J* = 7.8 Hz, 2H), 1.71 (s, 6H).

^{13}C NMR (125 MHz, $CDCl_3$, -40^{o}C) δ 206.4, 144.4, 143.4, 141.3, 141.1, 129.7, 128.9, 128.1, 127.1, 126.1, 125.4, 119.8, 118.9, 72.9, 29.0.

HRMS (ESI): 415.1690 [calculated mass for $C_{30}H_{23}O_2$ $(M+H)^+$ 415.1698]



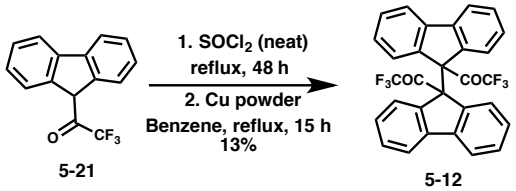
Compound 5-21: To an oven-dried 50 mL round-bottom flask with a stir bar was added fluorene (**5-2**) (1.0 g, 6.02 mmol) and potassium t-butoxide (1.01 g, 9.03 mmol, 1.5 eq). The solids were dissolved in THF (25 mL) and ethyl trifluoroacetate (1.1 mL, 9.03 mmol, 1.5 eq) was added dropwise. The reaction was heated at reflux under nitrogen for 20 h. The reaction mixture was poured into a saturated NH_4Cl solution and extracted three times with diethyl ether. The combined ether extracts were washed with water, dried over Na_2SO_4, and the solvent was removed *in vacuo* to give a yellow oil which slowly solidified. The compound was purified by silicia column chromatography (6:1 hexanes/ethyl acetate) to give compound **5-21** (1.002 g, 64%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 7.83 (d, J=7.5 Hz, 2H), 7.50 (m, 4H), 7.37 (t, J=7 Hz, 2H), 5.25 (s, 1H).

^{13}C NMR (125 MHz, $CDCl_3$) δ 190.4 (q, J=33.9 Hz), 142.8, 139.1, 129.6, 127.6, 125.7, 121.2, 116.2 (q, J=292.1 Hz), 57.3.

^{19}F NMR (470 MHz, $CDCl_3$, C_6F_6) δ: -78.30

HRMS (ESI): m/z 263.0691 [calculated mass for $C_{15}H_{10}OF_3$ $(M+H)^+$ 263.0684]



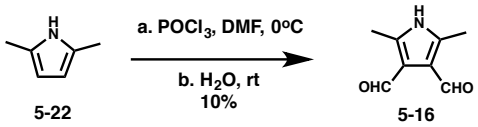
Compound 5-12: To a 50 mL round bottom flask containing compound **5-21** (500 mg, 1.91 mmol) was added thionyl chloride (20 mL). The reaction was heated at reflux for 48 h. The thionyl chloride was removed. ^{1}H NMR spectroscopy showed a compound consistent with the α-chloro trifluoromethyl ketone with partial conversion to the desired dimer. The crude material (426 mg) was added to a 10 mL round bottom flask containing copper powder (827 mg) and benzene (40 mL). The reaction was heated at reflux for 15 h after which the copper was filtered and the solvent evaporated. The product was purified by silica column chromatography (20:1 hexane/ethyl acetate) to give compound **5-12** (63.4 mg, 13%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.23 (bs, 2H), 7.55-7.49 (m, 8H), 7.10 (bs, 2H), 6.66 (bs, 2H), 5.76 (bs, 2H)

^{13}C NMR (125 MHz, $CDCl_3$) δ 144.5, 142.2, 140.3, 136.6, 130.2, 129.7, 129.4, 127.7, 126.4, 125.6, 120.3, 119.4, 115.7 (q, J = 294.5 Hz), 70.0.

^{19}F NMR (470 MHz, $CDCl_3$, C_6F_6) δ -74.80

HRMS (ESI) 523.1140 [calculated for $C_{30}H_{17}F_6O_2$ $(M+H)^+$ 523.1133]



5-22 5-16

Compound 5-16: A modified version of the procedure of Boudif and Momenteau was followed[12]. Briefly, $POCl_3$ (0.33 mL) was added dropwise to DMF (0.5 mL) at 0°C and stirred for 1 h. Next, 2,5-dimethyl-1H-pyrrole (**5-22**) (100 mg, 1.05 mmol, 1 eq) was added at room temperature and heated to 50°C for 2 h. The reaction was cooled to room temperature and slowly quenched by addition to H_2O. The reaction was extracted three times with EtOAc and washed with saturated aqueous NaCl. The solution was dried through Na_2SO_4, concentrated and purified by silica column chromatography to afford compound **S14** (16.4 mg, 10%) as a white solid.

^{1}H NMR (500 MHz, d_6-Acetone) δ 10.26 (s, 2H), 2.50 (s, 6H).

^{13}C NMR (125 MHz, d_6-Acetone) δ 186.5, 137.9, 120.4, 11.1.

HRMS (ESI) 152.0716 [calculated for $C_8H_{10}NO_2$ $(M+H)^+$ 152.0712]

5.8 References

1. Johnstone, R. W., Ruefli, A. A. & Lowe, S. W. Apoptosis: a link between cancer genetics and chemotherapy. *Cell* **108,** 153–164 (2002).
2. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144,** 646–674 (2011).
3. Carey, F. A. & Sundberg, R. J. *Advanced Organic Chemistry*. (Springer, 2007).
4. Bowden, K., El-Kaissi, F. A. & Ranson, R. J. Intramolecular catalysis. Part 5. The intramolecular Cannizzaro reaction of o -phthalaldehyde and [α,α′- 2 H 2]- o -phthalaldehyde. *Journal of the Chemical Society, Perkin Transactions 2* **0,** 2089–2092 (1990).
5. Leslie, B. J. & Hergenrother, P. J. Identification of the cellular targets of bioactive small organic molecules using affinity reagents. *Chem. Soc. Rev.* **37,** 1347–1360 (2008).
6. Curtin, D. Y., Kampmeier, J. A. & O'Connor, B. R. Nitrosation Reactions of Primary Vinylamines. Possible Divalent Carbon Intermediates'. *J. Am. Chem. Soc.* **87,** 863–873 (1965).
7. Palchaudhuri, R. & Hergenrother, P. J. Transcript profiling and RNA interference as tools to identify small molecule mechanisms and therapeutic potential. *ACS Chem. Biol.* **6,** 21–33 (2011).
8. Goldstein, J. C., Waterhouse, N. J., Juin, P., Evan, G. I. & Green, D. R. The coordinate release of cytochrome c during apoptosis is rapid, complete and kinetically invariant. *Nat. Cell Biol.* **2,** 156–162 (2000).
9. Matsuyama, S., Xu, Q., Velours, J. & Reed, J. C. The Mitochondrial F0F1-ATPase proton pump is required for function of the proapoptotic protein Bax in yeast and mammalian cells. *Mol. Cell* **1,** 327–336 (1998).
10. Borowiecki, P., Balter, S., Justyniak, I. & Ochal, Z. First chemoenzymatic synthesis of (*R*)- and (*S*)-1-(9*H*-fluoren-9-yl)ethanol. *Tetrahedron: Asymmetry* **24,** 1120–1126 (2013).
11. Greenhow, E. J., Harris, A. S. & White, E. N. The chemistry of fluorene. Part IV. Some new chloro- and nitro-derivatives. *J. Chem. Soc.* 3116–3121 (1954). doi:10.1039/jr9540003116
12. Boudif, A. & Momenteau, M. A new convergent method for porphyrin synthesis based on a '3 + 1' condensation. *J. Chem. Soc., Perkin Trans. 1* 1235 (1996). doi:10.1039/p19960001235